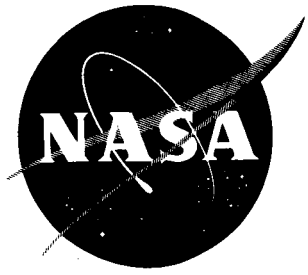
NASA TN D-305



TECHNICAL NOTE
D-305

ON THE DESIGN OF A HIGH-GAIN SATURATING CONTROL SYSTEM FOR USE AS AN ADAPTIVE AUTOPILOT

By John D. McLean and Stanley F. Schmidt

Ames Research Center
Moffett Field, Calif.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

WASHINGTON February 1960

TABLE OF CONTENTS

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

TECHNICAL NOTE D-305

ON THE DESIGN OF A HIGH-GAIN SATURATING CONTROL SYSTEM FOR USE AS AN ADAPTIVE AUTOPILOT

By John D. McLean and Stanley F. Schmidt

SUMMARY

A theoretical investigation is made of the use of a high-gain saturating control system for an adaptive autopilot. The results of the study show that such a system can control an aircraft over the entire flight envelope without requiring air data measurements. The capabilities and limitations of this type of control system are explored, and it is found that difficulties may be encountered at either high or low extremes of dynamic pressure. At high dynamic pressure, the limit cycle or "chatter" which is inherent in the system may produce objectionable or intolerable effects. Analytical methods are presented for predicting the frequency and amplitude of this chatter and its effect on the low-frequency performance. Methods are also given for reducing chatter amplitude and its undesirable effects on the low-frequency response. At low dynamic pressures, the performance is restricted by limiting and low aerodynamic gain. Attempts to make the system response time short for small inputs will result in a poorly damped response and, in some cases, instability for large inputs. It is shown that one way of compensating for this difficulty is to accept a sluggish response for low dynamic pressures; the closed-loop transfer function changes automatically with flight condition to take advantage of the higher aircraft capabilities at high dynamic pressures. It is also shown that the system can be designed to control unstable aircraft.

INTRODUCTION

Considerable interest has been focused recently on adaptive autopilots[1] as a means of avoiding the complexity and reliability difficulties of gain changers in the autopilot which heretofore have been used to cope with large variations in the aerodynamic characteristics of the aircraft.

[1]There is considerable controversy in the control systems field as to what constitutes an adaptive control system. A definition which has come into common usage in the case of autopilots is that an adaptive autopilot is one which will maintain its closed-loop transfer function invariant over the flight envelope without the use of air data measurements. In this report the closed-loop response is not required to be invariant but it may change in an acceptable manner.

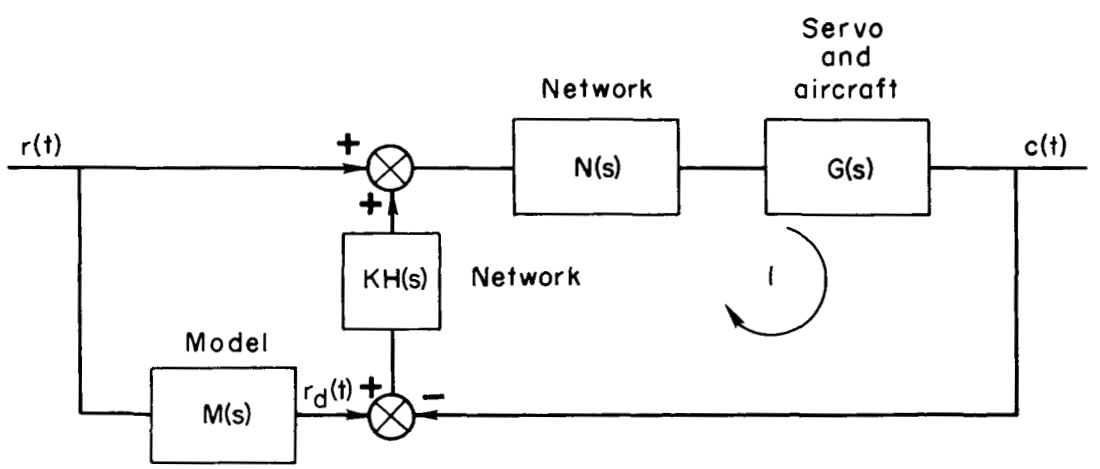
A large number of papers dealing with this subject can be found in the proceedings of a recent WADC symposium (ref. 1). One of these papers (ref. 1, p. 33) was presented by the authors. This report contains the material of that paper plus additional material which has since been derived.

The problem of designing an adaptive autopilot is essentially one of making the transfer function, relating the output and input of a feedback control system, independent of the aircraft transfer function. It is well known that this end can be accomplished by making the open loop gain very high. A very high gain system, however, generally has stability problems which heretofore have prevented its application to autopilots.

It is the purpose of this investigation to study the problems of a high gain system to determine the fundamental difficulties and practical solutions to these difficulties.

The scope of the investigation will be restricted to considering transfer functions in the longitudinal mode only and it will be assumed the aircraft equations of motion are linear. Many of the derived fundamental characteristics of such systems, however, can be extended to cover the other modes of the aircraft and nonlinear equations of motion.

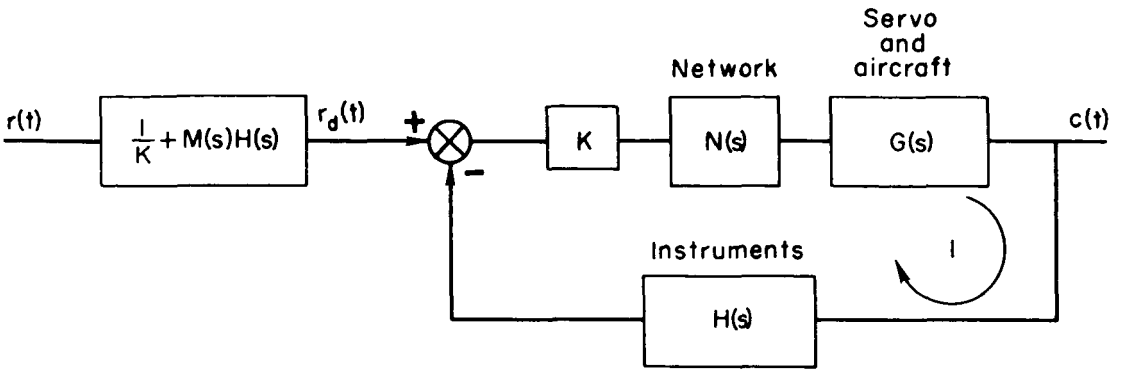
It is convenient to use the model concept in describing an adaptive autopilot. This description is aided by the block diagram of sketch (a).



Sketch (a)

Here, M(s) represents the transfer function of the ideal or model, that is, the desired transfer function relating the output to the input. The input signal is $r(t)$, and $c(t)$ represents the controlled output (e.g., pitch rate, normal acceleration, etc.). The same input is applied to both the model and the aircraft system, and if a difference or error signal exists between the desired response, $r_d(t)$, and the output, $c(t)$, then a corrective signal is applied through the network, KH(s), to make

the difference reduce to zero. This simple concept, although quite useful for explanatory purposes, is not complete. A requirement which must be met is that loop (1) of sketch (a) must be a stable loop. For purposes of understanding it is generally convenient to transform the block diagram of sketch (a) to the one of sketch (b).



Sketch (b)

This diagram is more useful for aircraft purposes since it allows H(s) to represent the measuring instruments which, as a result of the methods of measurement, must have certain dynamic properties. One can derive from sketch (b) the fact that one way to make the closed-loop transfer function, C/R, equal to the model transfer function, M(s), is to make the gain, K, infinite. Sketch (b) also shows that the transfer function

$$\frac{C}{R_d} = \frac{KN(s)G(s)}{1 + KN(s)H(s)G(s)} = \frac{1}{H(s)} \qquad \text{for} \quad K = \infty$$

Thus, a second alternative for the system is to make $R_d/R = 1$ and the model transfer function equal to $1/H(s)$. Again the system response will behave like the model. This last alternative is the one that will be concentrated on in this report. Other schemes can be found in reference 1.

As was mentioned with reference to sketch (a), loop (1) of sketch (a) or sketch (b) must be a stable loop. As will be shown, practical considerations always make it impossible for K to be infinite and the system to be stable. It will be shown, however, that if one installs a limiter or saturating amplifier in an appropriate place in the loop, the transfer function of C/R_d of sketch (b) can be made to approximate $1/H(s)$, the reciprocal of the feedback transfer function, and still be stable. The theory and method for analysis and design of such systems are derived in this report in the section entitled, "Analysis of a High-Gain Saturating System." An application of the theory and methods to a high-speed air-to-air missile is given in the section entitled, "A Missile Example."

NOTATION

a,b	arbitrarily chosen constants
A	normal acceleration, g
A_d	desired normal acceleration, g
B	limit level, deg/sec
$\bar{c}$	wing mean aerodynamic chord
c(t)	controlled quantity
C_c	chatter amplitude of the controlled quantity
C_L	lift coefficient
C_{L_α}	$\frac{\partial C_L}{\partial \alpha}$
C_{L_δ}	$\frac{\partial C_L}{\partial \delta}$
C_m	pitching-moment coefficient
C_{m_α}	$\frac{\partial C_m}{\partial \alpha}$
$C_{m_{\dot{\alpha}}}$	$\frac{\partial C_m}{\partial(\dot{\alpha}\bar{c}/2V)}$
C_{m_δ}	$\frac{\partial C_m}{\partial \delta}$
$C_{m_{\dot{\theta}}}$	$\frac{\partial C_m}{\partial(\dot{\theta}\bar{c}/2V)}$
C(s)	Laplace transform of controlled quantity
d	normalized dither amplitude
I_Y	pitching moment of inertia, slug-ft^2
K	gain
m	mass, slugs
m(t)	manipulated variable

P	magnitude of negative real pole
q	dynamic pressure, lb/sq ft
r(t)	reference quantity
R(s)	Laplace transform of reference quantity
s	Laplace transform variable
S	wing area, sq ft
t	time, sec
V	velocity, ft/sec
x(t)	saturated variable
y	normalized constant input
α	angle of attack, deg
β	unbalance in limit level
γ	flight path angle, deg
δ	control deflection, deg
ζ	damping ratio
θ	pitch angle, deg
τ	time constant, sec
φ	phase angle
ω	angular frequency, radians/sec

Subscripts

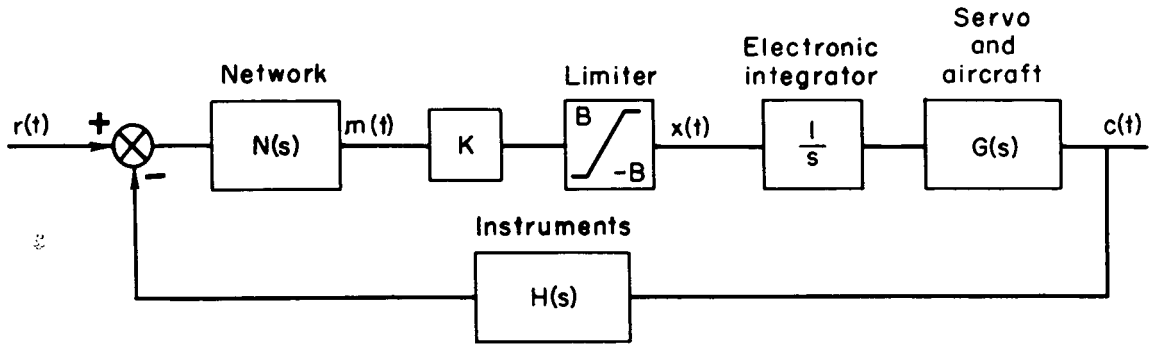
a	aerodynamic
c	crossover or chatter
d	desired response
f	filter
i	instruments

n	networks
s	servo
o	zeros in s plane
1,2	arbitrary parameters

ANALYSIS OF A HIGH-GAIN SATURATING SYSTEM

Basic Theory

The basic theory will first be used to demonstrate that a high-gain saturating control system corresponds to an adaptive autopilot.[2]



Sketch (c)

Sketch (c) is a block diagram typical of autopilots. The input command signal is r(t), and c(t) is output controlled motion (e.g., pitch rate, normal acceleration, etc.). The limiter has unity gain and is preceded by a gain K. Note that if K is infinite, the combination of the limiter and gain blocks is equivalent to an ideal relay since, then, $x(t) = B \operatorname{sgn} m(t)$. The term G(s) represents the transfer function of the servo-aircraft combination. For purposes of explanation, the servo transfer function will be assumed to be unity. An electronic integrator is located after the limiter. Since the servo transfer function is unity, x(t), which is bounded by the limiter, is the control-surface velocity.[3] Thus, we are dealing with an autopilot which is nonlinear as a result of limiting the control-surface rate.

[2]A high-gain saturating system has been used to obtain an adaptive autopilot by Minneapolis Honeywell (ref. 1, p. 123). The material presented here is felt to extend this previous work by presenting analytical methods for designing such systems.

[3]An equivalent system would result if the limiter were the hydraulic valve (with position limits) and the integrator were the approximate relation between valve position and control-surface deflection.

The system of sketch (c) is a saturated control system of a type studied previously (e.g., ref. 2). It has generally been shown that the characteristic response of this system will vary markedly with both the magnitude and frequency of the input. It has also been shown that a good visualization of the changes in the response characteristics with input magnitude and frequency can be obtained by drawing a root-locus diagram as a function of equivalent limiter gain.[4] The closed-loop poles are considered to be at positions on the loci which are determined by the equivalent gain of the limiter. Thus, for very large inputs the closed-loop poles are nearly at the open-loop position (i.e., the equivalent limiter gain is low). For very small inputs, the closed-loop pole locations can be determined by a method to be shown later.

To illustrate the use of the root-locus technique consider the following transfer functions of the blocks of sketch (c).

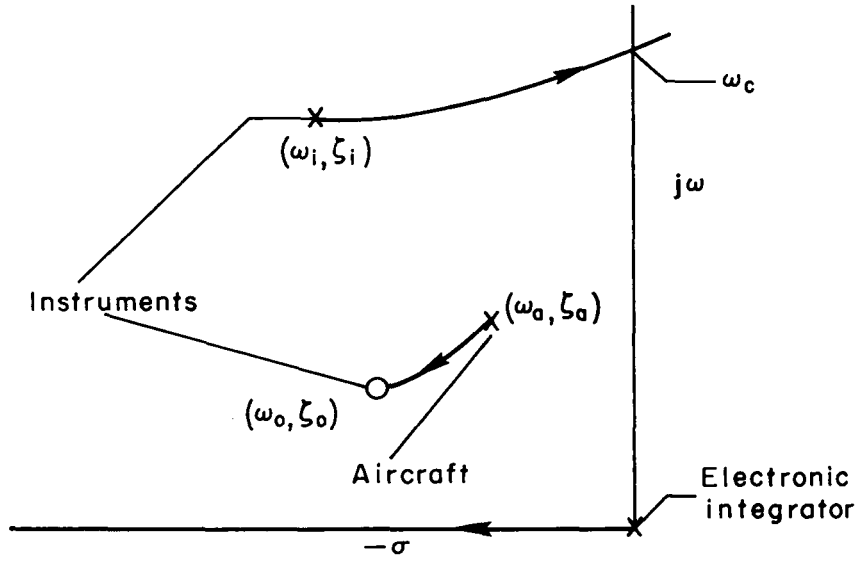
$$\left.\begin{aligned} G(s) &= \frac{K_a}{(s^2/\omega_a^2) + (2\zeta_a/\omega_a)s + 1} \\ H(s) &= \frac{(s_o^2/\omega_o^2) + (2\zeta_o/\omega_o)s + 1}{(s^2/\omega_i^2) + (2\zeta_i/\omega_i)s + 1} \\ N(s) &= 1 \end{aligned}\right\} \qquad (1)$$

For this example, G(s) is an approximate relationship between normal acceleration and control-surface deflection for a tail-controlled aircraft. The H(s) transfer function represents an approximate relationship for the sum of the outputs of a normal accelerometer, a pitch-rate gyro, and a pitching accelerometer as will be shown later. Thus, sketch (c) is the block diagram of a normal-acceleration autopilot. The characteristic equation for this system (without saturation) is readily derived from sketch (c). It is

$$1 + \frac{K}{s} G(s)H(s)N(s) = 0 \qquad (2)$$

[4]The limiter is treated as a device whose gain decreases as its input magnitude increases (see, e.g., ref. 2).

The loci as a function of the gain[5] K are shown in sketch (d) for one set of values of ω_a, ζ_a, ω_o, ζ_o, ω_i, and ζ_i. Sketch (d) illustrates that loci of the instrument poles (the complex conjugate is not shown) move into the right half plane for some finite value of K. Thus, if K is made infinite, the unsaturated system will be unstable. The system



Sketch (d)

is not unstable, however, since any divergence in the response would cause the input to the limiter to grow and consequently reduce its equivalent gain. As a result, the system with no inputs has a limit cycle (or chatter) of frequency ω_c. With reference to sketch (c), it can be seen that if ω_c is very high, the bounded output of the limiter, x(t), will be well filtered by the integrator, servo, and aerodynamics which follow it. Thus, the output is approximately zero when the input is zero.

This system is identical in many respects to off-on control systems investigated by Dr. Flügge-Lotz and her associates (refs. 3 and 4). The relay chatter in their experiments was attributed to relay imperfections, such as dead time. However, one can approximate these imperfections by a linear, second-order transfer function and then attribute the chatter (at ω_c) to the fact that the denominator of (K/s)N(s)G(s)H(s) is at least three orders greater than the numerator.[6] This fact can be understood by considering the asymptotic behavior of the root locus plots for such a system (e.g., see ref. 5).

---

[5]The actual K used in this study will be infinite; however, the root locus as a function of equivalent limiter gain will be the same as the one which is a function of the K in sketch (c).

[6]This condition (or approximation) is true for all practical systems when all high-order dynamics are considered (amplifier response, etc.).

For large step inputs or initial conditions, the response of this system may be obtained by phase space methods (ref. 4) or by the switch time method (for step inputs) of reference 2. For qualitative purposes, the closed-loop poles may be considered as moving back from the small-signal operating point (which is still to be determined) toward the open-loop pole positions as the input magnitude is increased. Thus, if the loci of the poles closest to the origin (the dominant poles) always stay in well-damped regions of the s plane as the equivalent gain is reduced, the system response will be reasonably well damped regardless of the input.

The small-signal operating point on the loci, that is, the poles and zeros of an equivalent linear system (valid for small or slowly changing input commands) can be found by determining the equivalent gain of the limiter. Note that gain at any point on the loci is the product of equivalent limiter gain and aerodynamic gain, K_a.

A 3 0 9

The equivalent gain of the limiter is found by treating its input as a constant or a low-frequency sinusoid, plus a constant-amplitude, high-frequency sine wave or "dither." If the limiter is treated in this manner, it can be shown (ref. 6) that for small inputs which are constants or sinusoids of lower frequency than the dither, the equivalent gain of the limiter is just half what it is for the dither input. A simple derivation of this fact, which is valid for small constant inputs is given in the appendix. Since the equivalent gain does not depend on the source of the dither, the chatter may be considered to be the dither. This simple relationship makes it very easy to calculate the positions of the closed-loop poles on the loci for small inputs on the basis of the following arguments:

1. The gain of the system (equivalent limiter gain times aerodynamic gain) for the dither signal must be the value of open-loop gain associated with the crossover point (ω_c of sketch (d)). This relationship is true since the open-loop gain must be high enough to sustain the limit cycle (or chatter) at ω_c.

2. Since the value of the gain for the low-frequency component of the input to the limiter is just half the value for the chatter, the equivalent linear system is found by moving back along the loci to the point corresponding to an open-loop gain of one-half that at ω_c.[7] Electronic simulation verifies that this simple process works exceptionally well as will be shown in the missile example section of this report.

---

[7]This analysis suggests that a system which measures the damping (ref. 1, p. 81) or frequency (ref. 1, p. 201) of the high-frequency mode and adjusts the open-loop gain so that the poles are in the left half plane, may be made to be less sensitive to parameter variations if it adjusts the open-loop gain to be more than one-half the crossover gain. It is doubtful, however, that the added insensitivity is worth the more complex hardware.

From this analysis one can say that if the gain associated with ω_c is very high (this will be true if ω_c is very high as is shown by conventional root-locus techniques, e.g., ref. 5), then all the closed-loop poles will be far away from the origin, except the two which have moved near the instrument zeros.

Thus, the resulting closed-loop transfer function is approximated by

$$\frac{C}{R} = \frac{1}{(s^2/\omega_o^2) + (2\zeta_o/\omega_o)s + 1} \tag{3}$$

Other investigators (ref. 3) have shown that in the chatter mode an on-off control system behaves like the system derived by setting the switching quantity equal to zero. For this example (neglecting instrument poles) an approximate switching quantity, m(t) of sketch (c), is

$$m(t) = r(t) - c(t) - \frac{1}{\omega_o^2}\ddot{c}(t) - \frac{2\zeta_o}{\omega_o}\dot{c}(t) \tag{4}$$

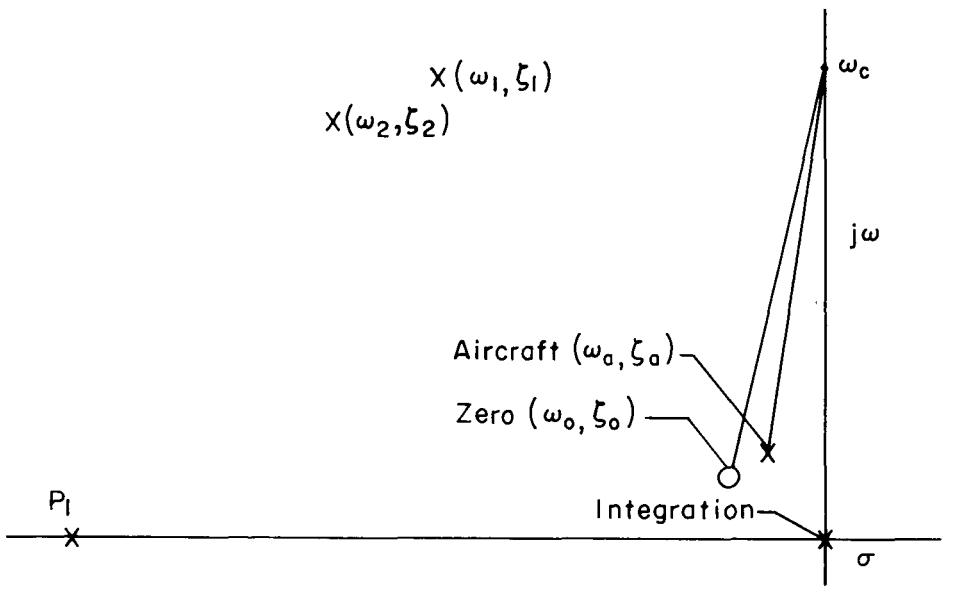
Setting m(t) equal to zero and using Laplace transform techniques one can derive equation (3) from equation (4). Thus, the root-locus arguments given here provide another analytical reason for the experimental results of reference 3.

The root-locus technique shown here is applicable for determining chatter frequencies and a qualitative idea of effects of input magnitude. Drawing root-locus diagrams for a large number of examples to determine preliminary designs, however, is quite a lengthy process. For this reason, a method for quickly determining the approximate chatter frequency and amplitude has been derived. This method along with two schemes for controlling the chatter amplitude, should it be excessive at certain flight conditions, is presented next.

Prediction of Chatter Frequency

In a practical autopilot of the high-gain saturating type, the chatter frequency must generally be much larger than the aircraft natural frequency or the natural frequency of the instrument feedback zeros. This condition must be met in order that the chatter at the output be reduced to a tolerable level. As a result of this condition, certain simplifications can be made such that the chatter frequency can be approximately calculated.

Consider the open-loop pole-zero plot shown in sketch (e):

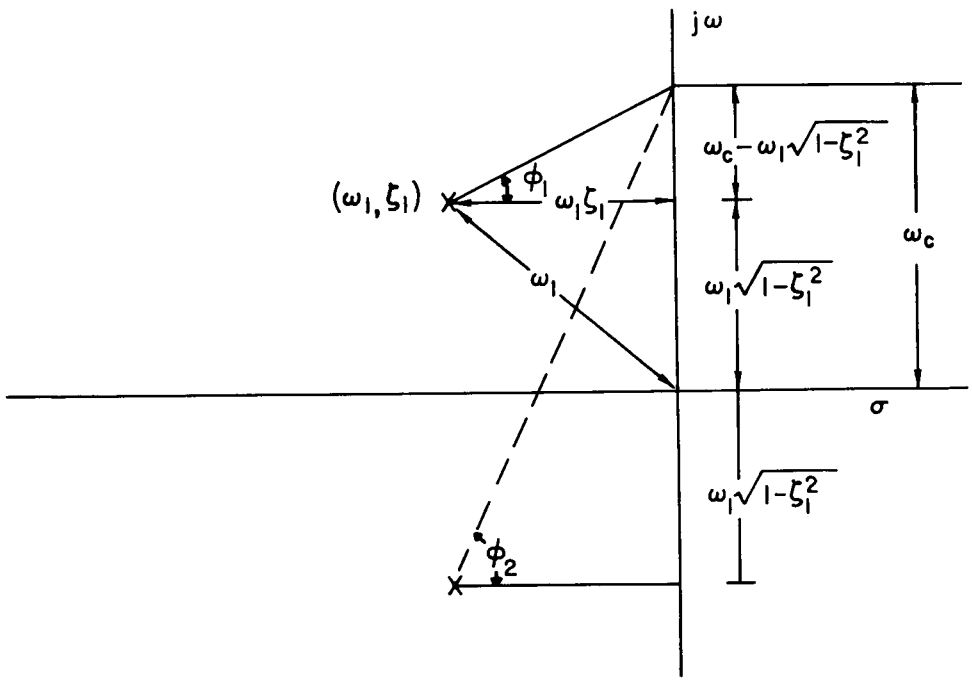


Sketch (e)

In this sketch, (ω_1,ζ_1), (ω_2,ζ_2), and (P_1) (and perhaps other poles not shown) represent the poles of instruments, servos, or networks of sketch (c). Their number and location depend upon the particular system under investigation.

As can be seen by the vectors drawn from the aircraft poles and the zeros to ω_c, the phase contributions at ω_c of the aircraft poles and the instrument zeros are nearly equal, but of opposite sign. Thus, for purposes of computing an approximate ω_c, they may be assumed to cancel each other. Since the phase shift at ω_c must be 180°, one can add up the phase angles of the other poles at ω_c to find the desired relationship between their location and the chatter frequency, ω_c. The process is simplified in this case since the integrator pole contributes 90° phase shift at all real frequencies. Thus, the sum of the phase shifts of the other poles (ω_1,ζ_1; ω_2,ζ_2; P_1, etc.) must be 90°.

To illustrate the method of computing ω_c as a function of the pole locations, consider the example in sketch (f). Here only two complex poles in addition to the aircraft and integrator poles are considered. This is therefore a fifth-order system.



Sketch (f)

From sketch (f) the following relationships are derived:

$$\tan \varphi_1 = \frac{\omega_c - \omega_1\sqrt{1-\zeta_1^2}}{\zeta_1\omega_1} \tag{5}$$

$$\tan \varphi_2 = \frac{\omega_c + \omega_1\sqrt{1-\zeta_1^2}}{\zeta_1\omega_1} \tag{6}$$

The trigonometric identity

$$\tan(\varphi_1 + \varphi_2) = \frac{\tan \varphi_1 + \tan \varphi_2}{1 - \tan \varphi_1 \tan \varphi_2} \tag{7}$$

is used to sum equations (5) and (6), giving

$$\tan(\varphi_1 + \varphi_2) = \frac{2\zeta_1\omega_1\omega_c}{\omega_1^2 - \omega_c^2} \tag{8}$$

Since $\tan 90^\circ = \infty$, and $\varphi_1 + \varphi_2$ must equal 90°, the denominator of equation (8) is set equal to zero, giving

$$\omega_c = \omega_1 \tag{9}$$

Equation (9) is the desired expression for this fifth-order example. For two real poles, $s = -P_1$, $s = -P_2$, it is easy to show that equation (9) holds if the following definitions are made:

$$\left.\begin{aligned} \omega_1^2 &= P_1 P_2 \\ 2\zeta_1\omega_1 &= P_1 + P_2 \end{aligned}\right\} \tag{10}$$

For higher order systems the same method may be used to obtain the equations relating ω_c and the locations of the poles. Table I gives these relationships for fourth-, fifth-, sixth-, and seventh-order systems. Note that for fourth order or lower, the chatter frequency does not exist since one can have infinite gain in these cases as can be readily shown by root-locus techniques.

In the solution for the seventh-order case, given in table I, a quadratic equation in ω_c^2 occurs. The frequency for 90° phase shift (ω_c) is the lowest frequency of the two solutions of the quadratic equation. The higher frequency solution will be for 270° phase shift. If realistic values of ζ_1 and ζ_2 are placed in the solution given, ω_c will be about $0.6\sqrt{\omega_1\omega_2}$. The quantity, $\sqrt{\omega_1\omega_2}$, is the geometric mean of the two frequencies. Consideration of higher-order cases than seventh shows that if the order is odd, the chatter frequency will be proportional to the geometric mean of the natural frequencies ω_1, ω_2, etc., of sketch (e) and must always be lower than the lowest frequency present.

The accuracy of the equations for chatter frequency given in table I depends on how nearly the phase lead of the instrument zeros and the phase lag of the aircraft poles cancel. With reference to sketch (e), it can be seen that if the aircraft poles are between the imaginary axis and the vectors drawn from the zeros to ω_c, then the poles contribute more phase lag than the lead contribution of the zeros. As a result, the actual chatter frequency will generally be somewhat lower than the predicted chatter frequency. If the opposite is true, the actual frequency is higher than the predicted frequency. This knowledge may serve to indicate when more accurate methods of computation are required.

Prediction of Chatter Amplitude

Once the chatter frequency is obtained the chatter amplitude can be readily calculated. For this calculation the output of the high-gain limiter is assumed to be a constant-amplitude square wave and only the fundamental component is used. The equation for chatter amplitude C_c (with reference to sketch (c)) is:

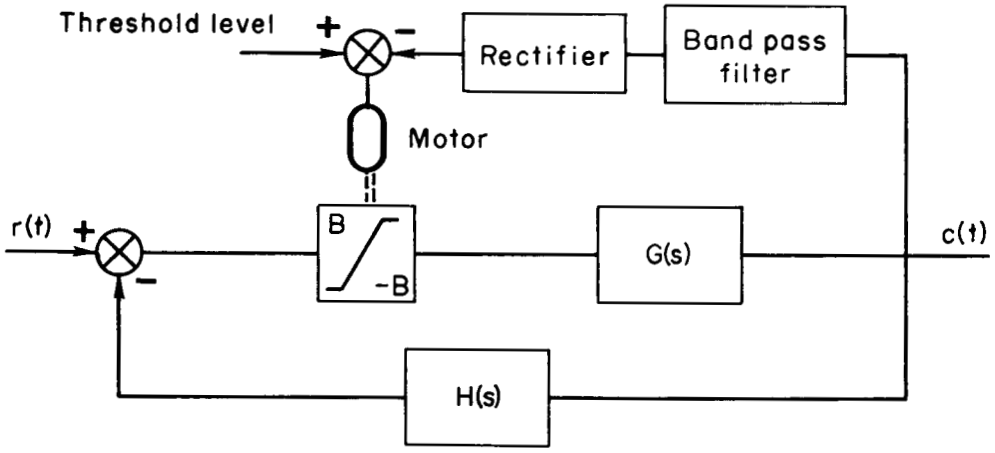
$$C_c = \left(\frac{4}{\pi} B\right)\left[\left|\frac{G(j\omega_c)}{j\omega_c}\right|\right] \qquad (11)$$

The quantity, $(4/\pi)B$, is the fundamental component of the limiter output. One may use linear methods, such as those used to derive equation (11), to derive equations for the chatter amplitude at any point in the loop. Equation (11) is very accurate since all the harmonic content of the square wave is well filtered by the integrator, servo, and aircraft.

Control of Chatter Amplitude

If the chatter amplitude is excessive for certain flight conditions and practical considerations make it impossible to increase the chatter frequency, then other means may be used to reduce the chatter. Two methods of accomplishing this objective are shown here. Both of these methods are mentioned in reference 7.

With reference to equation (11) it is seen that the chatter amplitude is directly proportional to B, the limit level. This suggests that one way of reducing chatter amplitude C_c is to reduce B if C_c exceeds a certain threshold level. A block diagram of a system for making this adjustment is shown in sketch (g).

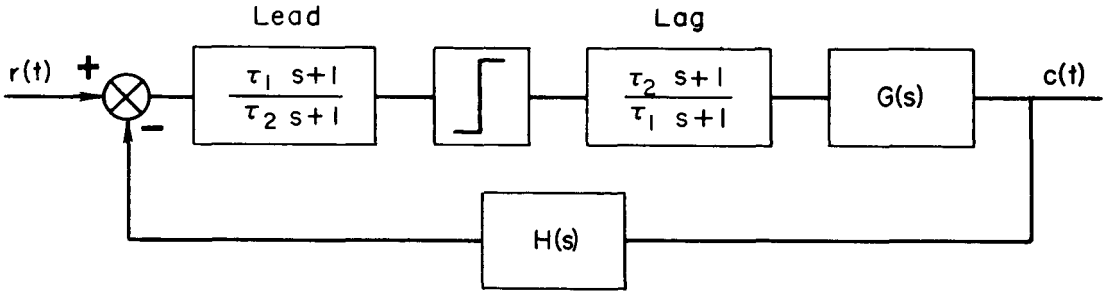


Sketch (g)

The output, c(t), is fed through a band-pass filter in order to separate the chatter from other output signals. The chatter is, in turn, rectified and compared to the threshold level. The motor turns to reduce the limit level (acting as an integrator) until the actual chatter is equal to the threshold value. Stops on the motor prevent B from ever exceeding some given maximum value should the error signal be positive. Certain precautions must be taken in the design of such a system to insure adequate stability of the compensating loop. The missile

example section of this report shows one satisfactory design which uses this approach. The success of the method depends on the fact that generally G(s) varies slowly, and therefore the compensating loop may be relatively sluggish. An obvious disadvantage is that considerable extra equipment is required.

A second method for controlling the chatter amplitude is based on the fact that the chatter amplitude at the output of the limiter is a constant. Since the chatter frequency is unaffected by networks which cancel each other (on a linear basis), one can place a lag network after the limiter and a lead network preceding the limiter. The lag network is adjusted to attenuate the chatter amplitude by the desired amount. Sketch (h) illustrates a block diagram of this approach.



Sketch (h)

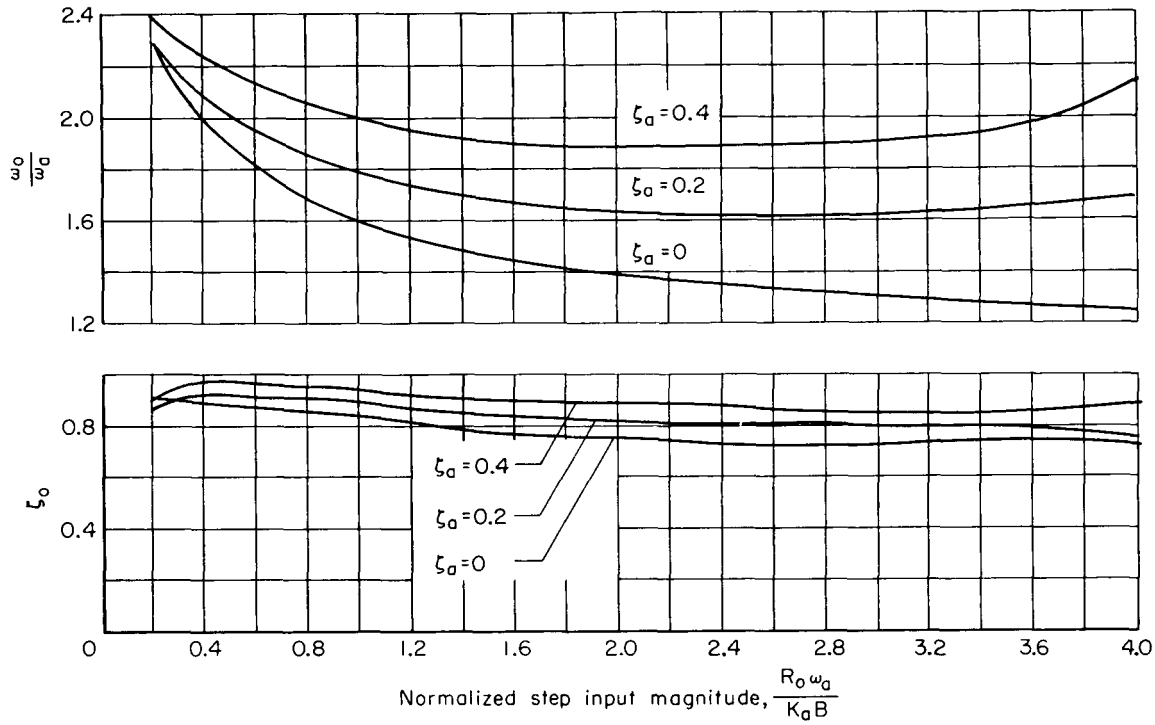
The advantage of this method is its simplicity. One disadvantage is that any high-frequency content of r(t) (e.g., noise) is amplified by the lead network. This high frequency has the same effect on the limiter as chatter; that is, it tends to reduce the equivalent gain of the limiter. Thus, apparently negligible noise on the input can be amplified by the lead network to such an extent that it reduces the equivalent gain of the system to the point where satisfactory performance cannot be obtained. As a matter of fact, noise on the input to any high-gain saturating system could render it useless for an adaptive system. This fact suggests that other techniques may have to be used in certain instances.

Effects of Zero Positions on the Response to Large Inputs

The report thus far has been concerned with the system behavior for small or slowly varying inputs. For these inputs the system is in the chatter region continuously and the system transfer function is nearly invariant with changes in the basic aircraft dynamics. For large input transients, such as steps, however, the system response can be oscillatory or unstable. The size of the step permitted in any particular system, before overshoot due to limiting occurs, can be obtained from

the switch time method of reference 2; however, control of the third-order system[8] represented by the integrator and G(s) of equations (1) has been studied extensively by Flügge-Lotz and Ishikawa (ref. 4). Their results can be used to illustrate the influence of the various aerodynamic parameters, the choice of zero positions, and the size step permitted before overshoot due to limiting occurs.

Sketch (i) shows one of the figures given in reference 4 replotted in a form more useful for our purposes.



Sketch (i)

The data shown in sketch (i) were obtained by adjusting the zero positions for each step-input magnitude for optimum step response. The optimum response for a third-order system requires two sign reversals of the saturated variable. It shows how the zeros should be adjusted as a function of the step-input magnitude. One point of significance is that the natural frequency of the zeros must be decreased as the step-input magnitude increases. This point agrees with the intuitive argument that if one of the output derivatives of the plant (airplane-servo combination) is limited, then the time of response of the system increases with the magnitude of the input. A second point of interest is that the damping of the zeros is relatively independent of the input magnitude over the range of inputs considered.

[8]For this analysis, the poles due to instruments, servos, etc., which cause the system to chatter are neglected.

It is necessary to understand that if the natural frequency of the zeros is too high for any particular step-input magnitude, the system will overshoot as a result of limiting. If the natural frequency is too low, a chatter region exists. Thus, if one desires to set the zeros for invariant autopilot response over the flight envelope and the maximum step-input magnitude is fixed, then one must choose the desired response at or lower than the optimum ω_c given for the maximum R_o[9] at the worst flight condition, that is, the flight condition when the aerodynamic gain and natural frequency are lowest. Otherwise, steps at certain flight conditions can result in oscillatory or unstable response. For aircraft, the worst flight condition generally occurs at the lowest dynamic pressure (the highest altitude and lowest velocity point of the flight envelope). At higher dynamic pressures faster response is possible, but cannot be obtained unless the zeros are shifted with flight condition. This method has been studied for a normal-acceleration autopilot and is considered later in this report.

Sketch (i) can be used to determine the highest frequency zero positions for the lowest dynamic pressure. For example, assume that for this flight condition $\omega_a = 1$, $K_a = 0.1$, $R_o = 2g$, $\zeta_a = 0$, and $B = 30^o/sec$. The value of the abscissa is then,

$$\frac{R_o \omega_a}{K_a B} = \frac{2}{3} = 0.667$$

The ordinate reading is $\omega_o = 1.77$ and $\zeta_o = 0.86$. For other autopilots (e.g., pitch rate) data such as presented in sketch (i) are not available. One, therefore, would resort to simulation studies in order to determine the zero positions. If the positions are to be invariant, however, one must always choose them for the largest step input at the worst flight condition.

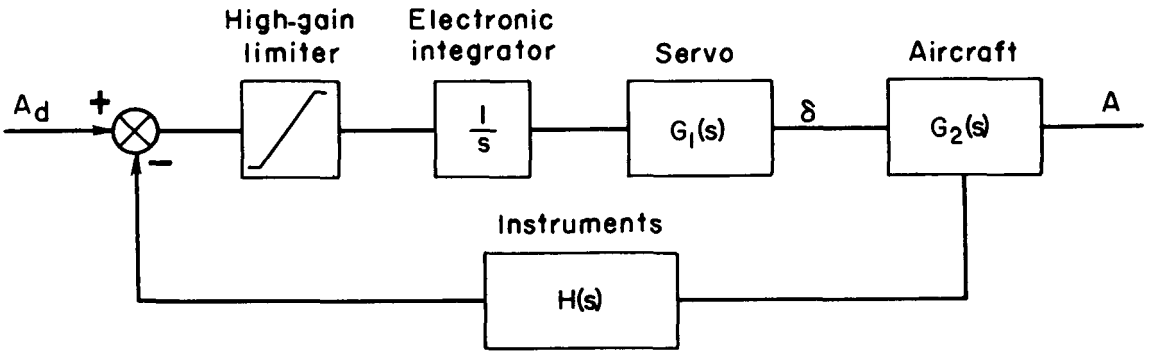
A MISSILE EXAMPLE

The purpose of working an example is to verify the theory and design methods proposed, and to determine what limitations, if any, exist in the use of a high-gain saturating control system as an adaptive autopilot. As an example, a hypothetical high-speed air-to-air missile, capable of attacking supersonic bombers from sea level to over 100,000 feet altitude, has been chosen. This example was chosen to include very marked changes in the aerodynamic characteristics over the flight envelope and to include natural frequencies of the airframe which are higher (compared to instrument frequencies) than for a conventional airplane. Filtering of the chatter to an acceptable output level, therefore, may be a problem. If success is attained here, this type system certainly should be feasible for airplanes which have a considerably smaller flight envelope.

[9] This is true only for a linear feedback, H(s). Nonlinear feedbacks can give optimum response for each magnitude of input but their use is beyond the scope of this report.

The autopilot to be designed will be of the normal-acceleration type; that is, an input voltage will command a normal acceleration. Such autopilots are in general use for many missiles and they have the distinct advantage of providing a simple means (limiting the input voltage) for limiting structural load (or g). Analog computer simulation will be used to verify the design approach.

The block diagram of the system is shown in sketch (j).



Sketch (j)

This block diagram is practically identical to sketch (c), so no description is considered necessary.

System Description

Aerodynamics.- The simplified aircraft equations of motion are as follows:

$$mV\dot{\gamma} = qS(C_{L_\alpha}\alpha + C_{L_\delta}\delta) \tag{12}$$

$$I_Y\ddot{\theta} = qS\bar{c}\left(C_{m_\alpha}\alpha + C_{m_\delta}\delta + \frac{\bar{c}C_{m_{\dot{\alpha}}}\dot{\alpha}}{2V} + \frac{\bar{c}C_{m_{\dot{\theta}}}\dot{\theta}}{2V}\right) \tag{13}$$

Conventional Laplace transform techniques can then be used to derive the following transfer functions:

$$\frac{A(g)}{\delta(\text{deg})} = \frac{V}{1845}\,\frac{K_a[(s^2/\omega_{\dot{\gamma}}^2) + (2\zeta_{\dot{\gamma}}/\omega_{\dot{\gamma}})s + 1]}{(s^2/\omega_a^2) + (2\zeta_a/\omega_a)s + 1} \tag{14}$$

$$\frac{\dot{\theta}(\text{deg/sec})}{A(g)} = \frac{1845}{V}\,\frac{\tau_{\dot{\theta}}s + 1}{(s^2/\omega_{\dot{\gamma}}^2) + (2\zeta_{\dot{\gamma}}/\omega_{\dot{\gamma}})s + 1} \tag{15}$$

The constant, 1845, in equation (14) is the conversion factor from feet per second squared per radian to g per degree. The coefficients of the transfer functions are as follows:

$$K_a = - \frac{C_{m_\delta} C_{L_\alpha} - C_{m_\alpha} C_{L_\delta}}{(C_{L_\alpha} \bar{c} C_{m_{\dot{\theta}}}/2V) + (mV/qS) C_{m_\alpha}}$$

$$\omega_{\dot{\gamma}}^2 = \frac{qS\bar{c}}{I_Y} \frac{C_{L_\alpha} C_{m_\delta} - C_{L_\delta} C_{m_\alpha}}{C_{L_\delta}}$$

$$\omega_a^2 = \frac{(qS)^2 \bar{c}}{I_Y} \left(- \frac{C_{m_\alpha}}{qS} - \frac{\bar{c} C_{m_{\dot{\theta}}} C_{L_\alpha}}{2mV^2} \right)$$

$$\zeta_{\dot{\gamma}} = - \frac{qS\bar{c}^2}{4\omega_{\dot{\gamma}} V I_Y} (C_{m_{\dot{\alpha}}} + C_{m_{\dot{\theta}}})$$

$$\zeta_a = \frac{1}{2\omega_a} \left[\frac{qS}{mV} C_{L_\alpha} - \frac{qS\bar{c}^2 (C_{m_{\dot{\alpha}}} + C_{m_{\dot{\theta}}})}{2VI_Y} \right]$$

$$\tau_{\dot{\theta}} = \frac{mV C_{m_\delta} - (qS\bar{c}/2V) C_{L_\delta} C_{m_{\dot{\alpha}}}}{qS(C_{m_\delta} C_{L_\alpha} - C_{m_\alpha} C_{L_\delta})}$$

An examination of representative values of the various aerodynamic coefficients will show that for a tail-controlled aircraft, the following simplifications can usually be made:

$$\frac{A}{\delta} \approx \frac{V}{1845} \frac{K_a}{(s^2/\omega_a^2) + (2\zeta_a/\omega_a)s + 1} \tag{16}$$

$$\frac{\dot{\theta}}{A} \approx \frac{1845}{V} (\tau_{\dot{\theta}} s + 1) \tag{17}$$

$$K_a \approx - \frac{qS}{mV} \frac{C_{m_\delta} C_{L_\alpha}}{C_{m_\alpha}} \tag{18}$$

$$\omega_a^2 \approx - \frac{qS\bar{c} C_{m_\alpha}}{I_Y} \tag{19}$$

$$\tau_{\dot{\theta}} \approx \frac{mV}{qSC_{L_\alpha}} \tag{20}$$

Although the terms neglected in the A/δ transfer function are quite small, the numerator of the transfer function for a tail-controlled aircraft has a positive real root, which can produce instability in the closed-loop system. This root can always be eliminated from the overall characteristic equation of the system by proper use of pitch rate and appropriate derivatives in the feedback path.

Three flight conditions were chosen: two to represent the extremes of the flight envelope and one intermediate case. The parameter values for these conditions are summarized in table II. This example allows studies of an autopilot where gain, K_a, variations of the aircraft are 80 to 1, natural frequency, ω_a, variations are of about 7 to 1, and damping ratio, ζ_a, variations are from 0.30 to 0.

Instrumentation.- As has been shown, the instrument feedback transfer function, $H(s)$, must contain two zeros, and the natural frequency and damping ratio of these zeros (for a high chatter frequency) are approximately equal to the natural frequency and damping ratio of the desired (or model) transfer function.

In order to provide instrumentation which will give two zeros, two approaches which are theoretically possible are as follows:

1. Use a normal accelerometer to measure the normal acceleration, $c(t)$, and feed its output through a second-order lead network.

2. Sum the output of a normal accelerometer through a first-order lead network with the output of another first-order lead network fed from a pitch-rate gyro.

For the first approach, the natural frequency and damping of the zeros is set by the network and therefore does not change with flight condition. Normal accelerometer noise may make this approach impractical since the network must have a gain which increases with frequency. In another type autopilot (e.g., pitch attitude) where instrumentation is available to measure the necessary derivatives of $c(t)$, an invariant zero position feedback transfer function is possible without the lead network. The problems which may be encountered with this example can therefore be expected in these other types of autopilot.

The second approach is more practical and it is usually employed without the lead network in a conventional normal-acceleration autopilot. In this approach the zero positions change as the flight condition changes. The position change can be partially controlled in the design, so that at high dynamic pressure the natural frequency of the zeros is higher than at low dynamic pressure. Thus, the response of the system using this approach will be faster at high dynamic pressure where faster response is possible. The response will become slower with reduced dynamic pressure where attempts to obtain a fast response may result in instability. This system thus does not provide the usual adaptive autopilot; however, for certain missile (and perhaps airplane) applications its features may be more desirable than invariant response.

The variation in zero positions with this second approach is most conveniently found by assuming the networks are capable of perfect differentiation. The sum of the two network outputs will then be given by

$$y = a_o\ddot{\theta} + b_o\dot{\theta} + c_o\dot{A} + d_oA \tag{21}$$

The use of equations (16) and (17) and Laplace transform techniques for finding the required derivatives allows one to put equation (21) in the following form

$$\frac{Y(s)}{A} = \left(\frac{1845}{V}\,\tau_{\dot{\theta}}a_o\right)s^2 + \left[\frac{1845}{V}\,(\tau_{\dot{\theta}}b_o + a_o) + c_o\right]s + \left(\frac{1845}{V}\,b_o + d_o\right) \tag{22}$$

The quantity $(1845/V)b_o + d_o$ will be chosen as unity at some mid-velocity so that the over-all system (sketch (j)) will have unity gain at this velocity. The gain of this system will thus change with flight condition. In conventional autopilots the gain change is eliminated by means of a low-frequency differentiating circuit (often called a washout circuit in this application) in the pitch-rate feedback path. For purposes of simplicity d_o will be adjusted so the gain is unity even though in practice the washout circuit would be used.

The natural frequency, ω_o, and damping ratio, ζ_o, of the zeros can be determined from equation (22) giving (for $(1845/V)b_o + d_o = 1$)

$$\omega_o = \frac{1}{\sqrt{(1845/V)\tau_{\dot{\theta}}a_o}} \tag{23}$$

$$\zeta_o = \frac{1}{2\sqrt{(1845/V)\tau_{\dot{\theta}}a_o}}\left[\frac{1845}{V}\,(\tau_{\dot{\theta}}b_o + a_o) + c_o\right] \tag{24}$$

Substitution of $\tau_{\dot{\theta}}$ (eq. (20)) in equation (23) gives

$$\omega_o \approx \frac{1}{\sqrt{(1845/V)(mV/qSC_{L_\alpha})a_o}} = \frac{1}{\sqrt{(1845/I_Y)\left\{[mC_{m_\alpha}\bar{c}/(qS\bar{c}/I_Y)C_{m_\alpha}C_{L_\alpha}]\right\}a_o}} \tag{25}$$

which by use of equation (19) gives:

$$\omega_o \approx \frac{\omega_a}{\sqrt{-1845\ a_o(mC_{m_\alpha}\bar{c}/I_YC_{L_\alpha})}} \tag{26}$$

Since equations (19) and (20) are approximate, equation (26) shows ω_o is approximately proportional to ω_a.

Equation (24) can be put in the form

$$\zeta_o = \frac{1}{2}\left[\frac{b_o}{a_o\omega_o} + \omega_o\left(\frac{1845}{V}\,a_o + c_o\right)\right] \tag{27}$$

The damping ratio, ζ_o, therefore changes as ω_o changes; however, it will be shown later that by appropriate selection of the constants b_o and c_o this change (for the example missile) can be small enough to give a reasonable system.

The dynamics of the instrumentation will be assumed to be approximated by a first-order time lag of time constant $\tau_i = 0.0125$ second. This representation is permissible as long as the over-all system is at least fifth order so that chatter must exist. As will be shown later, as far as chatter frequency is concerned, this time constant provides the same phase shift as a second-order system with a natural frequency of 22.8 cycles per second and damping ratio of one half.

Servo.- The dynamics of the servo are approximated by a first-order lag of time constant $\tau_s = 0.025$ second. As will be shown, the same chatter frequency would exist if the servo were second order with a natural frequency of 11.4 cps and damping ratio of one half.

The limiter will be set so that the maximum control surface velocity is 30° per second.

Control-surface position limiting has not been considered in this study and in some cases may necessitate alterations to the system. When position limiting occurs the system response is simply that of the aircraft alone, and in manned vehicles or ones where stability is a problem it may be necessary to reserve some control-surface deflection for stability augmentation. However, this problem has been successfully coped with in conventional autopilots and should be no more difficult to deal with in the system under study here.

System Design and Simulation

Specifications.- The following specifications were selected for the resultant closed-loop normal-acceleration autopilot.

1. A step input of at least 2.5 g for the worst flight condition (condition 1 of table II) should have a good step response; that is, it should not be unstable or extremely oscillatory. For the other two flight conditions it would be desirable to have good step response for inputs as high as 20 g, the assumed structural load limit.

2. The desired transfer function, the reciprocal of H(s) of sketch (j) shall be either

(a) Fixed by the zero positions selected to meet specification (1)

(b) Allowed to vary in accordance with equations (26) and (27)

In either case, however, the damping ratio, ζ_o, given by equation (27) shall be in the range $0.4 < \zeta_o < 0.9$. This variation in ζ_o would be reasonable for most normal-acceleration autopilots.

3. The chatter amplitude at any flight condition shall not exceed 0.2 g (peak to peak).

Selection of zero positions.- The data presented in sketch (i) can be used to select the zero positions for the optimum response to a given magnitude of step input to this system if one neglects instrument and servo time lags. The step magnitude chosen is 2.5 g and aerodynamic data for condition 1 of table II.

The abscissa of sketch (i) is

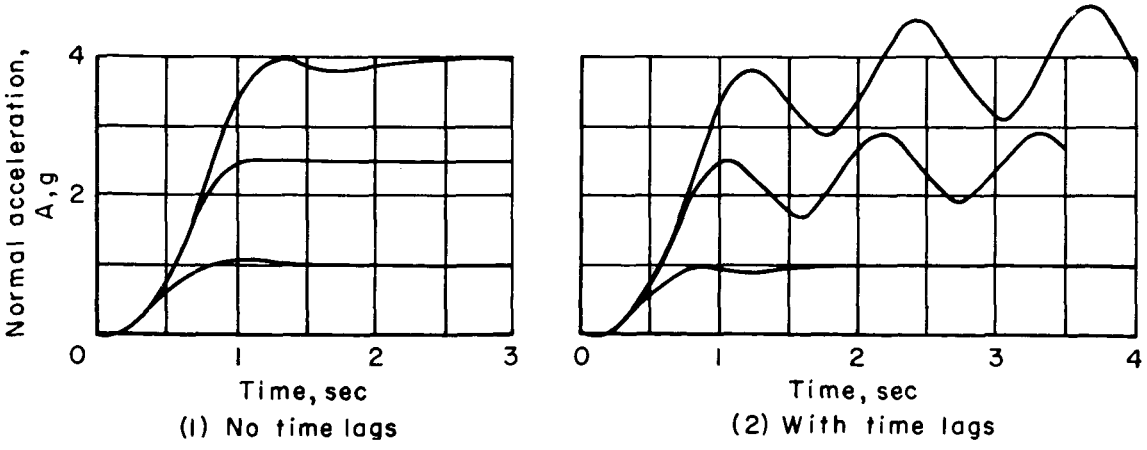
$$\frac{R_o \omega_a}{K_a B} = \frac{2.5(3.6)}{0.11(30)} = 2.73$$

This gives the zero positions described by

$$\omega_o = 4.75$$

$$\zeta_o = 0.72$$

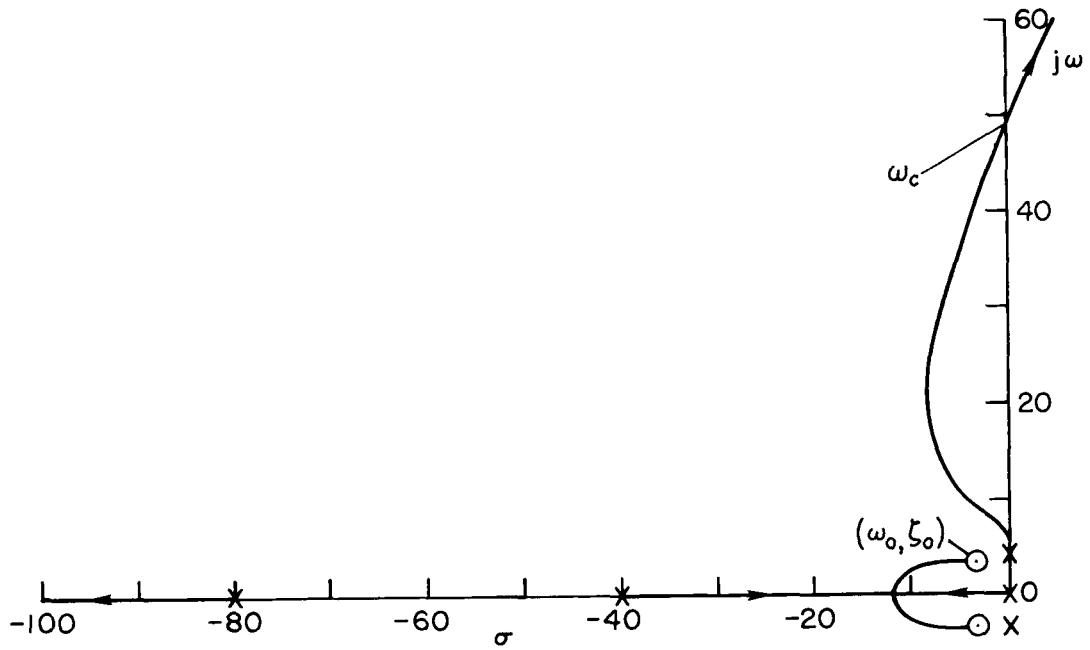
In order to check the validity of the assumption of neglecting the servo time lags, the system was simulated on an analog computer and steps of various magnitudes were applied for no instrument or servo lags and for the lags assumed for this example. The responses are shown in sketch (k).



Sketch (k)

Without time lags an optimum (no overshoot) response is obtained for a 2.5 g step. The introduction of time lags for this system, however, causes the response to become unstable for a 4 g input and to be completely unacceptable for 2.5 g. This result illustrates the fact that time lags of the magnitudes given certainly are not to be neglected in selecting the zero positions. The reasons for the oscillatory and unstable performance are best explained by drawing the root locus as a function of equivalent limiter gain.

This is shown in sketch (l). As can be seen from sketch (l) the locus from the aircraft pole is in the right half plane for low values of limiter gain. Thus, large inputs (or noise signals) can shock the system into an unstable mode (see ref. 2). This fact suggests a better way of

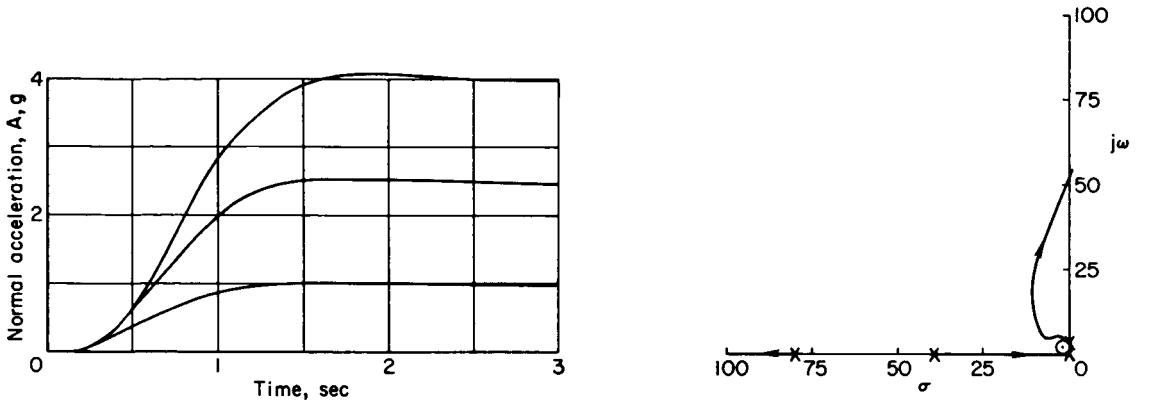


Sketch (l)

choosing zero positions in order to prevent any instabilities in the system due to inputs. The method is to choose the maximum ω_o for the desired value of damping, such that the angle of departure from the complex pole (for this example) is greater than 90°. There is a certain amount of arbitrariness to this choice, however; the values chosen for this example are

$$\left.\begin{aligned} \omega_o &= 2.89 \\ \zeta_o &= 0.679 \end{aligned}\right\} \tag{28}$$

The step responses and root loci for these zero positions are given in sketch (m). As can be seen from the root loci, no instability can result in this system for low values of equivalent limiter gain. The step



Sketch (m)

responses are well damped for step inputs as high as 4 g and are somewhat slower than the optimum given by sketch (k). The zeros given by equations (28) will therefore be used in the fixed zero example to be studied here.

For the case with moving zeros, the values of a_o, b_o, and c_o of equations (23) and (24) are chosen so that equations (28) are satisfied for the worst flight condition.

Equation (23) gives

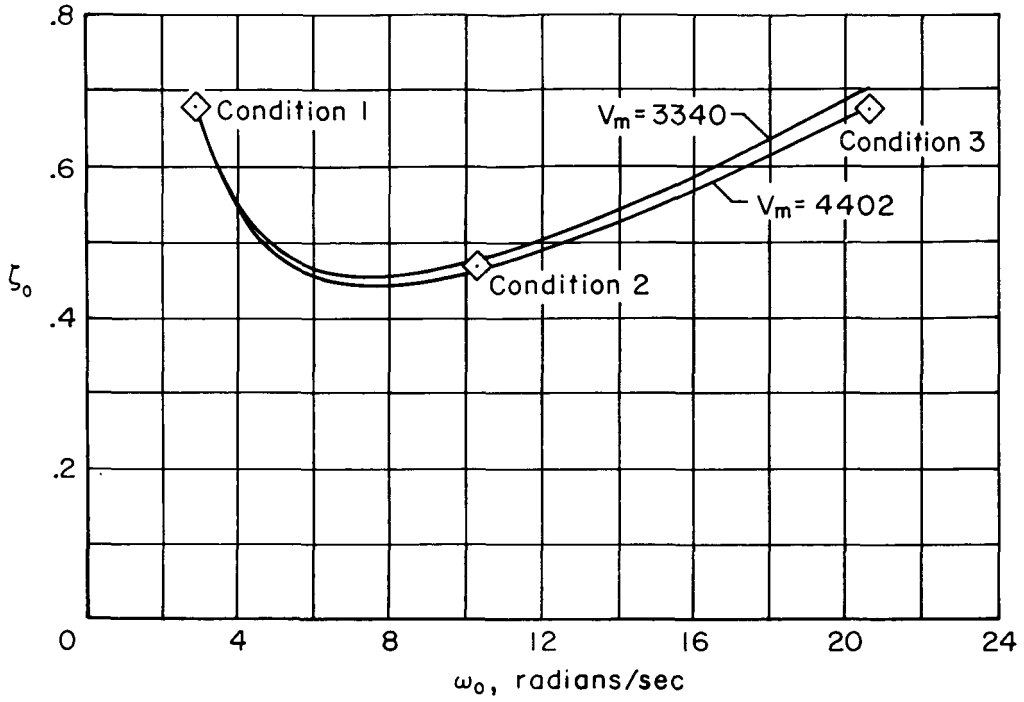
$$a_o = 0.0181 \tag{29}$$

Equation (27) shows that ζ_o cannot be chosen as a constant, that is, it must vary with ω_o and V. Since both b_o and c_o can be selected, some control of the variation is possible. The numbers selected here are

$$\left.\begin{aligned} b_o &= 0.0618 \\ c_o &= 0.0502 \end{aligned}\right\} \tag{30}$$

The variation of ζ_o with ω_o for V = 3340 and V = 4402 is shown in

sketch (n). Although the damping ratio of the desired response is less than one-half for certain flight conditions, this system will be considered as satisfactory.



Sketch (n)

Chatter amplitude prediction and control.- The approximate chatter frequency can be determined from table I. This frequency is invariant with flight condition and is equal to

$$\omega_c = \sqrt{\frac{1}{\tau_s}\frac{1}{\tau_i}} = 56.6 \text{ radians/sec} \qquad (31)$$

The formula of table I for the seventh-order case shows that ω_c would be the same if we had assumed second-order characteristics for the servo and instruments and had $\zeta_s = \zeta_i = 0.5$, $f_i = 22.8$ cps, $f_s = 11.4$ cps. These numbers for natural frequency are reasonable, but somewhat conservative in terms of currently available instruments and high performance hydraulic servos.

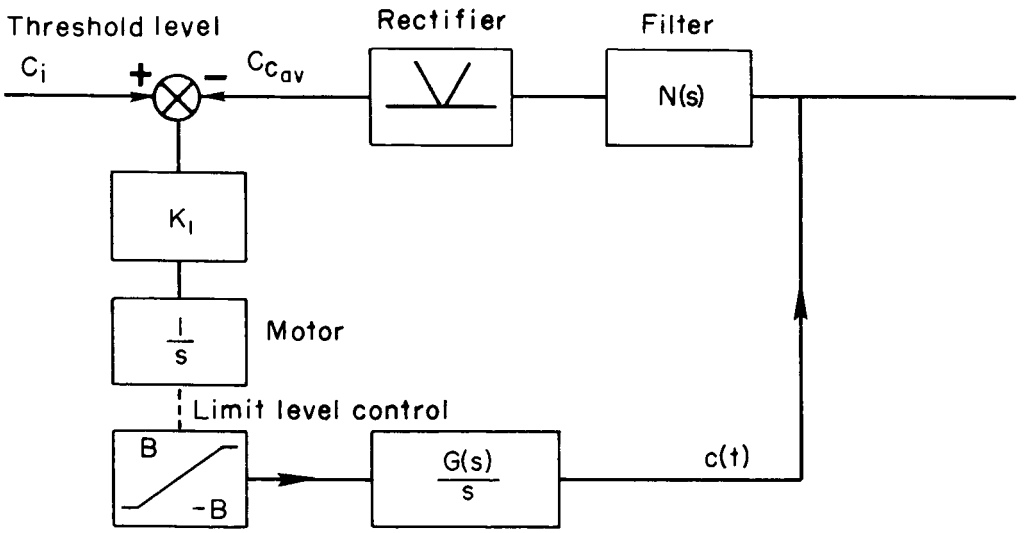
To illustrate the accuracy of the approximation for this example, the chatter frequency was determined by the root-locus method and by simulation. In this particular case, the chatter frequency given by equation (31) is not high compared to the aircraft natural frequency or natural frequency of the zeros for all cases; thus, the accuracy of the prediction is questionable. The results are given in table III and are reasonably good. The disagreement between the results of the simulation

and root-locus method can be attributed to both simulation inaccuracy and unaccounted-for time lags of the analog computing elements. The inaccuracies, however, definitely suggest that the actual chatter frequency in any particular system can only be determined by mockups including hardware.

Once the chatter frequency is known, the chatter amplitude can be determined by equation (11). The chatter amplitude obtained by use of the formula and root-loci computed chatter frequencies as well as the measured chatter amplitude are given in table IV. The chatter amplitude based on chatter frequency computed by formula and by the root locus are not in good agreement. The latter amplitude, however, is reasonably close to the measured value and the difference is attributed to the differences in chatter frequency.

The data of table IV show that the peak-to-peak value of g exceeds the specifications (0.2 peak to peak) for conditions 3 and 5. This indicates that either faster instruments and servos must be used (to increase chatter frequency thereby reducing the amplitude) or that some means of controlling chatter amplitude must be employed. The latter method is investigated here using the schemes of sketches (g) and (h).

If the method represented by sketch (g) is used, then one must design the compensating loop for adequate performance. This loop is shown in sketch (o):



Sketch (o)

It will be assumed that G(s) changes rather slowly so that a time constant of approximately 1 second for this compensating loop is sufficiently fast. The problem is to determine K_1 and N(s) of sketch (o) for the desired performance.

Equation (11) shows that $C_c = [(4/\pi)B]|G(j\omega_c)/j\omega_c|$. Thus, if one assumes that $|N(j\omega_c)| = 1$ and that any dynamics associated with the transfer function relationship $C_{c_{av}}/B$ are negligible, then the constant relationship between $C_{c_{av}}$ and B is

$$\frac{C_{c_{av}}}{B} = \frac{2}{\pi}\frac{4}{\pi}\left|\frac{G(j\omega_c)}{j\omega_c}\right| \tag{32}$$

The $2/\pi$ factor of equation (32) is used to obtain the average value of the assumed sine wave.

The time constant, τ_c, of the loop is equal to the reciprocal of the gain around the loop or

$$\tau_c = \frac{\pi^2}{8}\left|\frac{j\omega_c}{G(j\omega_c)}\right|\frac{1}{K_1} \tag{33}$$

Since $|j\omega_c/G(j\omega_c)|$ varies with flight condition, K_1 must be made to vary with flight condition if τ_c is to remain constant. It is probable that in many cases K_1 could be selected at some constant value and the change in τ_c would not be objectionable. There is, however, a simple means of compensation if this is not the case. This is shown as follows: Assume that the loop is functioning correctly and the motor is not on its stops. Then

$$C_i = C_{c_{av}} = \frac{8}{\pi^2}\left|\frac{G(j\omega_c)}{j\omega_c}\right|B \tag{34}$$

or

$$B = \frac{\pi^2}{8}\frac{C_i}{|G(j\omega_c)/j\omega_c|} \tag{35}$$

where C_i is a constant. Thus, the limit level, or shaft position of the motor, can be used to obtain a gain which decreases as $|G(j\omega_c)/j\omega_c|$ increases. The constant, K_1, of sketch (o) can be replaced by a variable by using a potentiometer with a variable gain, K_s, on the motor shaft. Since K_s must be less than unity, the motor will be assumed to have a gain, K_m, which can have any selected value.

For design purposes assume $B = 30^\circ$ per second when $K_s = 1.0$, or $K_s = B/30$. From equation (35) is obtained

$$K_s = \frac{\pi^2}{8}\frac{C_i}{[|G(j\omega_c)/j\omega_c|]30} \tag{36}$$

Since $K_sK_m = K_1$, the time constant of the loop is (from eq. (33))

A 3 0 9

$$\tau_c = \frac{30}{C_i K_m} \tag{37}$$

For $\tau_c = 1$ and $C_i = (2/\pi)(0.1)g$

$$K_m = 300\left(\frac{\pi}{2}\right) = 472 \tag{38}$$

The network transfer function $N(s)$ must be selected next. This network must have the characteristics of a band pass filter with zero gain for frequencies less than ω_c since it is undesirable to pass any signals resulting from inputs to the closed-loop system. The filter should, in general, cut off above ω_c in order to eliminate noise of the type resulting from measurements and should have reasonable damping for transient performance due to disturbances such as gusts. These requirements generally imply that the use of a filter of the form

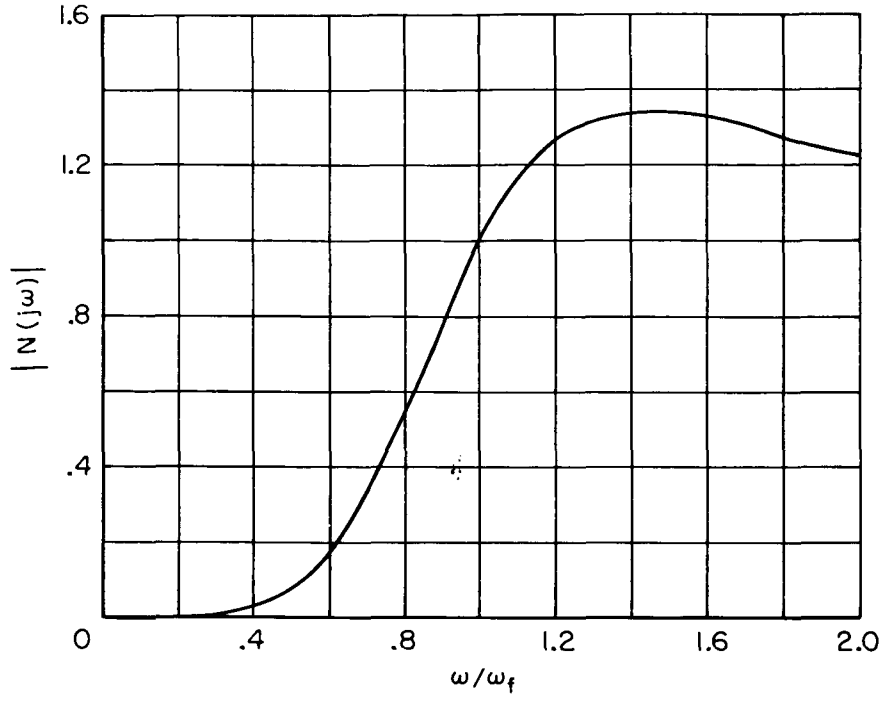
$$N(s) = K\left\{\frac{[s^n/(\omega_f)^n][(s^2/\omega_1^2)+1][(s^2/\omega_2^2)+1]}{[(s^2/\omega_f^2)+(2\zeta/\omega_f)s+1]^{n+4+x}}\right\} \tag{39}$$

could result in reasonable performance. The quantities x and ω_2 in equation (39) are determined by the high-frequency cutoff characteristics desired; n and ω_1 are determined for the low-frequency characteristics desired. The design of an optimum filter for this application is beyond the scope of this investigation. The filter which was used to demonstrate the principle of this type system is given by

$$N(s) = \frac{s^4/\omega_f^4}{[(s^2/\omega_f^2)+(s/\omega_f)+1]^2} \tag{40}$$

It should be emphasized that n of equation (39) should be as high as permissible (from practical construction standpoints) in order to prevent signals due to input responses from affecting the compensating loop. The value $n = 4$ was found to be satisfactory for this example, and it was also found that there was no need to attenuate frequencies higher than ω_c since the computer introduced little noise. In general, it is believed however, that the success of such systems which detect and adjust themselves by means of a test signal (chatter in this case) depends greatly on the use of a relatively complex network.

The frequency response of the network used is shown in sketch (p). This sketch and the data of table III show that ω_f should be chosen at about 40 radians per second in order that the time constant τ_c and



Sketch (p)

chatter amplitude C_c be within 10 percent of the design values. For the chatter frequency variation calculated (table III) the network gain shown in sketch (p) is about 1.26. This value is accounted for in the system.

The design of the system of sketch (h) for controlling chatter amplitude is simply a matter of choosing the constants τ_1 and τ_2. In practice, high-order networks may be necessary; however, only first-order networks will be considered here.

The chatter amplitudes for cases 3 and 5 of table IV are seen to exceed the specified values. For the example, only the fixed zero case will be considered so that only filtering for case 3 need be determined. A desired attenuation of 1/8 is indicated by the data in table IV obtained from the approximate ω_c. Since the approximation formula provides a conservative estimate of ω_c, the use of the data from this formula provides a margin of safety of about 2 to 1. As can be seen from table IV, a factor of 1/4 would satisfy case 3; however, other points of the flight envelope could have a higher chatter amplitude.

If it is assumed that $\tau_2 = 0.1\ \tau_1$, then the reduction of the network at ω_c is

$$\frac{1}{8} = \sqrt{\frac{0.01\ \tau_1^2\omega_c^2 + 1}{\tau_1^2\omega_c^2 + 1}} \tag{41}$$

Assuming $\omega_c = 56.6$, one solves equation (41) obtaining

$$\tau_1 = 0.234 \tag{42}$$

and since $\tau_2 = 0.1\ \tau_1$

$$\tau_2 = 0.0234 \tag{43}$$

Different values than those given in equations (42) and (43) can be obtained by different ratios of τ_2/τ_1; however, this ratio must be smaller than the chatter amplitude reduction desired. The method of selection also gives the smallest value of τ_1 permissible for the τ_2/τ_1 ratio assumed. This result is considered desirable from the standpoint of having the least effect on the response for large step inputs.

Equivalent linear system for small inputs.- Analysis has shown that the high-gain saturated control system behaves in a linear fashion for inputs which are small enough to have little effect on the inherent limit cycle. In particular it has been shown that the equivalent gain of the limiter for signal inputs is just half what it is for the chatter. Therefore, a root-locus graph of the system as a function of equivalent limiter gain allows one to pick off the pole positions of the equivalent linear system. The graphs for the five cases are shown in figure 1. The heavy dots on the loci indicate the positions for the gain being just half the gain where the aircraft poles cross into the right half plane. These heavy dot positions can therefore be read off and an equivalent linear system transfer function derived.

For case 1 the dominant poles (those closest to the origin) are very close to the feedback zeros and the other poles are quite a large distance away. This simply implies that the transfer function is closely approximated by a second-order transfer function whose denominator is the feedback zeros. For the other cases this is not true. The heavy dots for cases 2 and 3 are on the real axis. Furthermore, for cases 2 and 3 the system is really dominant first order. Thus, with the instrument and servo dynamics assumed it is impossible to have a high enough gain of the system to obtain an equivalent linear system whose approximate transfer function is the reciprocal of the instrument feedback zeros. The only way this could be accomplished is by use of higher performance equipment. The approximate transfer functions indicated by the heavy dots, however, are not bad in the sense of being oscillatory. Instead, the system response is more sluggish than desired and may or may not be undesirable in the particular over-all task (e.g., homing or beam riding).

Response for large inputs.- As has been shown here and previously (ref. 2), the root-locus graph can be used to obtain a qualitative picture of performance for inputs large enough to reduce the equivalent limiter gain. As a result of the choice of zeros for case 1, it is seen from the loci for all cases that no locus crosses into the right half plane as the equivalent gain is reduced to zero. Thus, as the input is increased, no kind of instability will result. As a matter of fact, the response will become dominant first order for all cases for very large inputs because of the pole at the origin. The step response for all examples will therefore be characterized by very little or no overshoot.

Simulation of the design.- The system response to steps has been checked by means of simulation. Prior to presenting the data it is desirable to review the proposed designs. Three examples have been worked out. These are:

1. Fixed zeros and a network used for chatter amplitude control.

2. Fixed zeros and a limit level controller used for chatter amplitude control.

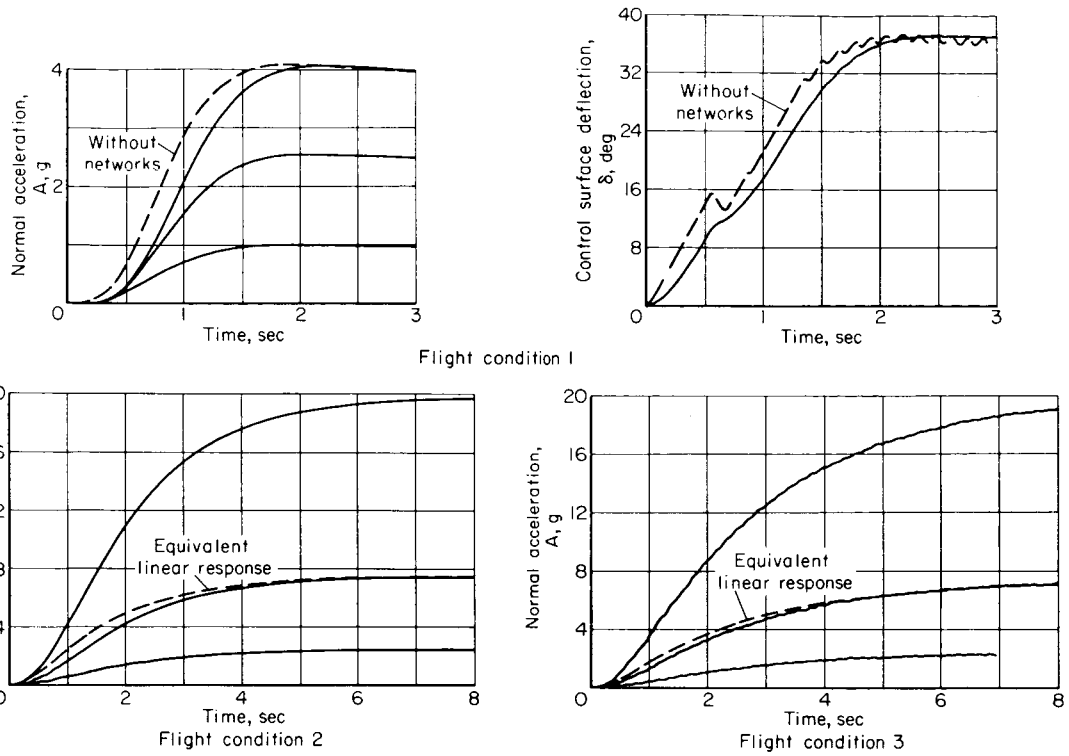
3. Moving zeros and a limit level controller used for chatter amplitude control.

The block diagrams for the three examples showing all the constants are given in figure 2. The transfer function coefficients of the missile for the three flight conditions studied are given in table II.

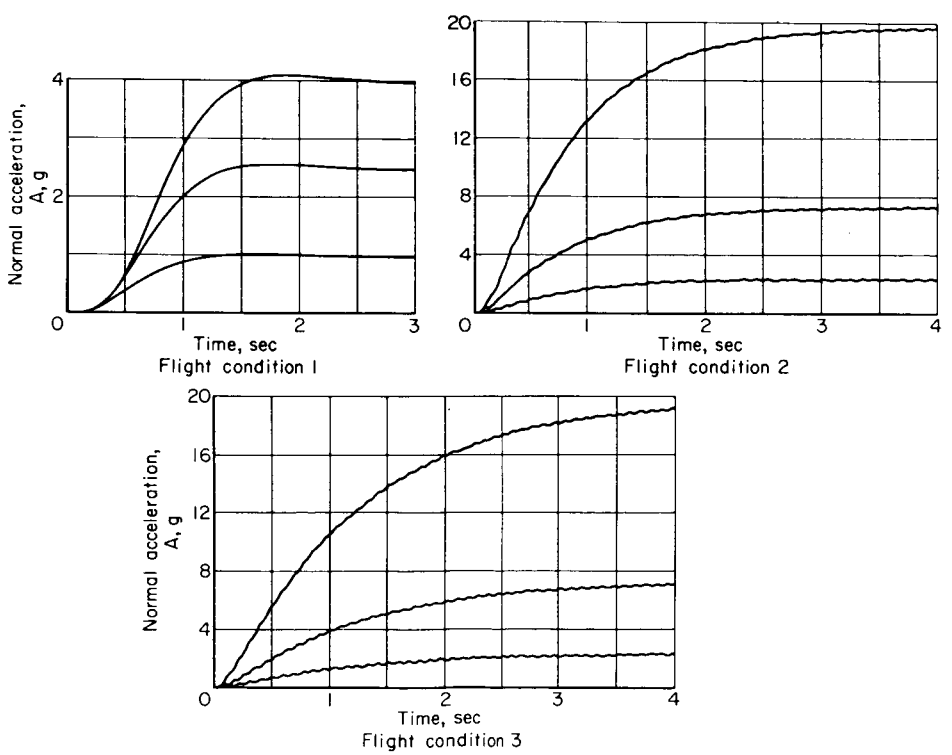
The step responses for example 1 for the three flight conditions are shown in sketch (q). The dotted line shown (for condition 1) indicates the response without the networks and shows that the networks have the effect of slowing down the response for step inputs. This is true for all flight conditions wherein a large amount of limiting takes place during the transient. The control surface position is shown for flight condition 1 (for the largest input) and illustrates the chatter present without the networks and the filtering of chatter produced by the network.

For conditions 2 and 3, the response is seen to be quite sluggish; however, no type of instability is present for inputs as high as 20 g, a result which agrees with the theory presented earlier. The response of the equivalent linear system (obtained from the heavy dots of fig. 1) is shown for conditions 2 and 3 which illustrates that even for these large transients, the system response agrees reasonably well with the equivalent linear system.

The step responses for example 2 are given in sketch (r). The limit-level controller has no effect on the transient performance of this example; that is, its only effect was to reduce the chatter amplitude for flight condition 3.

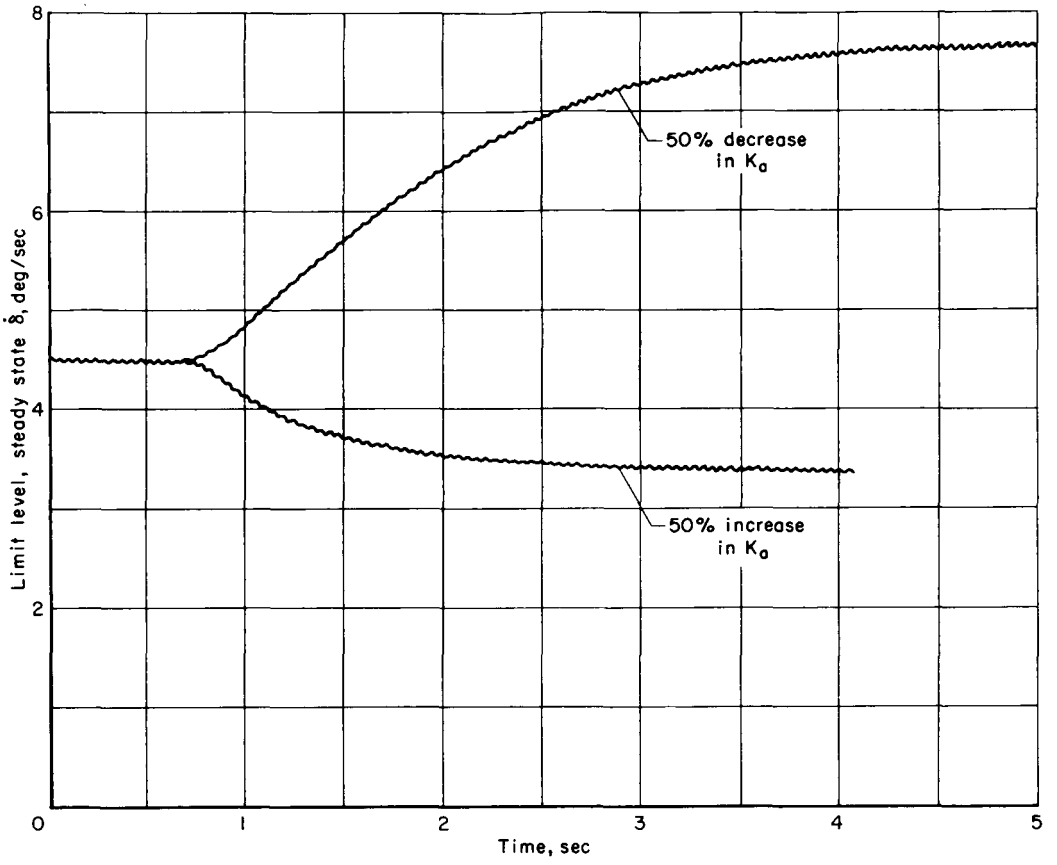


Sketch (q)



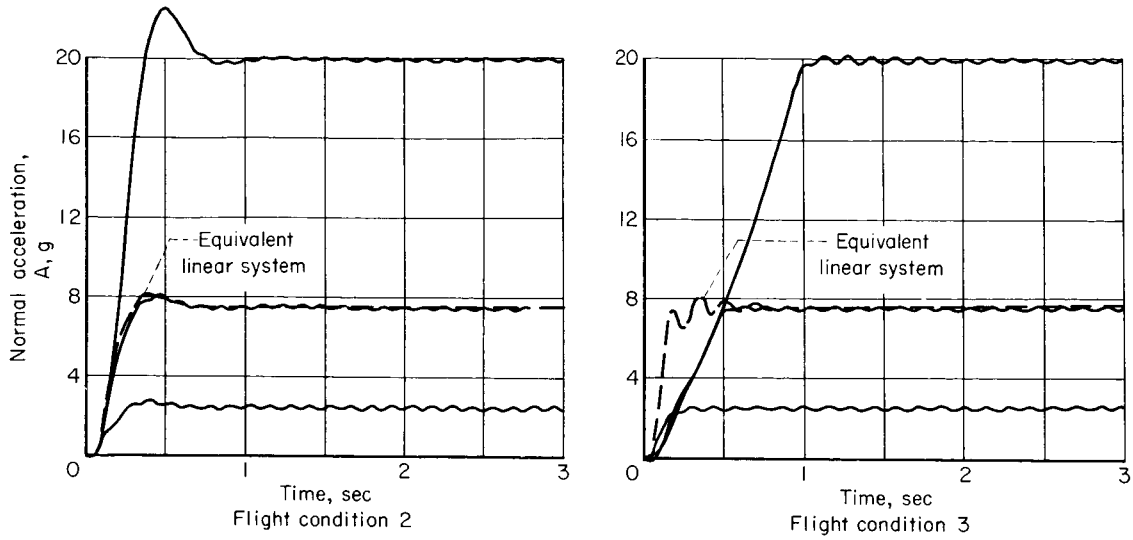
Sketch (r)

To illustrate the time response of the limit-level controller, 50-percent step changes in K_a were run and the limit level recorded. These data are presented in sketch (s). The response corresponds closely to the design value of 1 second time constant. These data were run for example 2, condition 3.



Sketch (s)

The step responses for example 3 are presented in sketch (t). Only flight conditions 2 and 3 are presented since for condition 1 the response is identical to that in sketch (r). The limit-level controller has no effect on the step response for condition 2 since the chatter amplitude is small, but it does have a significant effect on the step response for condition 3. As a result of the high chatter amplitude (without control) for this example (see table IV), the maximum rate of control-surface deflection, that is, the limit level, must be reduced from 30° per second to approximately 1.43° per second in order to reduce the chatter amplitude to the specified level. This large reduction in limit level results in a response which is saturated during most of the transient for large inputs. It is interesting to note that as a result of this excessive limiting the response is slower for condition 3 than for condition 2.



Sketch (t)

The step response for 7.5 g input for the linear system obtained from figure 1 is shown dotted for comparison. The comparison for condition 2 is quite good; however, for condition 3, only for inputs less than about 2 g does the system behave in a linear fashion. The response for this example is certainly faster than for example 2 for all but condition 1; and therefore using moving zeros is a reasonable way of obtaining faster response when the aircraft is capable of providing this increase.

Discussion of results.- For the missile and hardware chosen in these examples it is not possible to have the gain of the system high enough to provide a response which is the reciprocal of the feedback zeros. The reason for this is simply that the chatter frequency is too low for all but condition 1. If one had higher performance hardware, which gives higher chatter frequencies, not only would the response have corresponded closely to the zeros but use of chatter amplitude control would not have been necessary.

For the large variation in parameters chosen, the change in response is not tremendous and it is probable that for many applications this magnitude of change would be permissible.

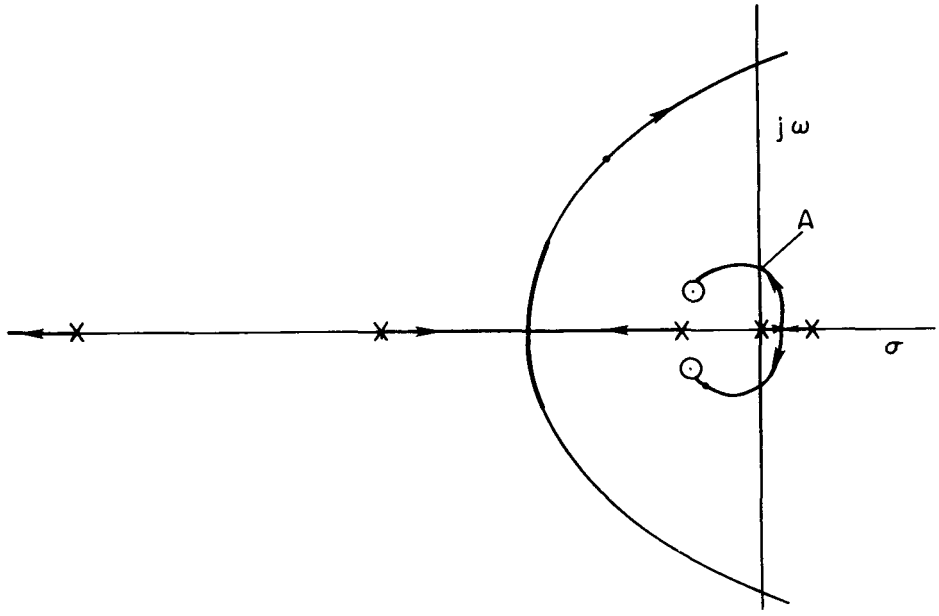
Of the two types of chatter amplitude control, the scheme using network is the most desirable from the simplicity standpoint. It does have the disadvantage, however, of slowing down the response during transients where a large amount of saturation is present. The design of the limit level controller appears to be straightforward. The only item that appears to be critical is the order of the network used to separate the chatter from the signals required for the aircraft to follow input commands. The limit-level controller has the disadvantage of being more complicated and the limit level may be driven down (to reduce chatter) to points where the response for large inputs becomes slow as a result

of saturation. This fact suggests that perhaps additional inputs to the limit-level controller (e.g., magnitude of the error) might be used to make the response faster for large inputs or large errors. It was beyond the scope of this investigation to consider the additional refinements which lead to nonlinear controllers.

Other Considerations

The particular airframe chosen for example purposes had a very high natural frequency for condition 3 which resulted in a high amplitude chatter. It is very probable that a reduction in static margin would have simplified the design problems. Shifts in static margin, however, may cause the airframe to be statically unstable under certain conditions. This imposes the question of whether a high-gain saturating control system can control an unstable airframe and what pitfalls, if any, exist with such a system. This section is devoted to consideration of this subject.

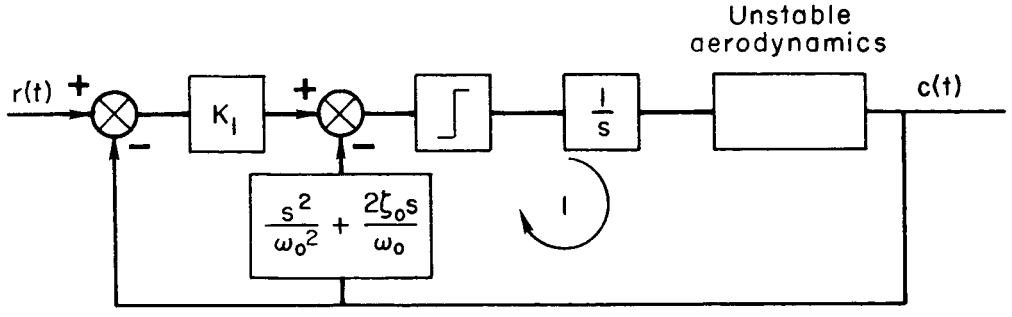
Assume the aircraft controlled in the block diagram of sketch (j) is statically unstable. The root locus of such a system as a function of equivalent limiter gain could be as shown in sketch (u).



Sketch (u)

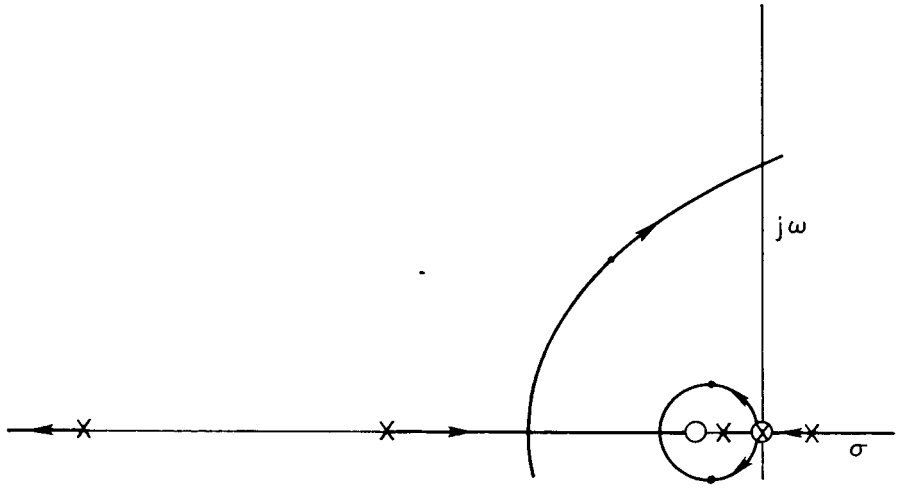
Assume that the heavy dots on the loci represent the pole locations for the equivalent linear system. It is quite evident that the equivalent gain of the limiter must never be allowed to fall below the value corresponding to point A (approximately) on the loci, otherwise an unstable

response will occur. This means that the system may be unstable for certain inputs. If these inputs are within the range of interest, then some means must be introduced to prevent the instability. This generally implies the use of nonlinear controllers, and if first reversal time data for this type plant were available, the switch time method of reference 2 could be used. Since these data are not available, use will be made of the root-locus method of considering the zeros to shift as a function of error in order to show what type of nonlinear function would prevent the instability. Consider the system of sketch (v). If K_1 is unity for



Sketch (v)

small errors, the root locus will be as shown in sketch (u). Suppose K_1 is actually a limiter which has unity gain for small inputs and saturates for some input level. If the output of the box is saturated, the root loci of loop 1 for this condition can be considered as shown in sketch (w).



Sketch (w)

This loop may have the equivalent linear system indicated by the heavy dots on the loci. If the limiter (in place of K_1) is set sufficiently low, then the equivalent limiter gain can never be driven low enough by inputs (error in this case) to cause system instability. Thus,

the introduction of a simple limiter can prevent system instability due to inputs. Gusts, of course, could still cause system instability and the only preventative for this is to be sure the control-surface rate limit, B, is high enough to prevent instability for any gust inputs the system is likely to encounter.

The nonlinearity introduced in place of K_1 is not the optimum, however; the form (i.e., saturation type) is suggested, and analog computer simulation could be used to find the shape which is best for the particular airframe under consideration.

CONCLUSIONS

A theoretical investigation of the use of a high-gain saturating control system for an adaptive autopilot has been made. From the results of this investigation, the following conclusions may be drawn:

1. A high-frequency chatter must exist in such a system, and if it can be made to have a very small amplitude, the autopilot can successfully control the aircraft over very large ranges of flight conditions.

2. Reduction of the chatter amplitude to a tolerable value depends on having the natural frequencies of the hardware components as high as possible compared to that of the aircraft. Methods may be devised to reduce the chatter amplitude below its natural value, but they will not alleviate the harmful effects of chatter on the low-frequency character of the response.

3. The speed of response at low dynamic pressures will be restricted because of limiting in combination with low aerodynamic gain, and attempts to force faster response will result in poor stability.

4. If the fastest response obtainable at low dynamic pressure is not satisfactory over the entire flight envelope, it may be necessary to shift the zeros with chatter amplitude, error magnitude, or both. In addition, it may be desired to change the limit level with chatter amplitude. The use of devices to accomplish these changes results in the possibility of unstable gain adjustment loops, but a combination of high gain and self-adjustment should result in a more versatile system.

5. There is a need for further study of systems which adjust zero positions and limit level in combination with the high-gain saturating system. In particular, it would be desirable to determine what nonlinear

compensation schemes, such as used in NASA TN D-20, might be added to improve the system characteristics. In addition, further studies which show what type nonlinear system is required for controlling unstable airframes would be very desirable.

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, Calif., Oct. 19, 1959

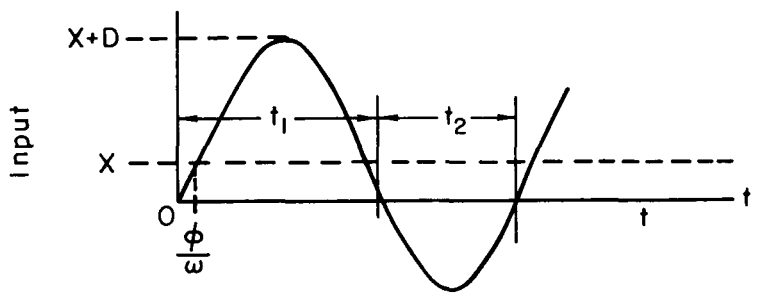
APPENDIX A

DERIVATION OF LIMITER GAIN IN THE PRESENCE OF DITHER

Consider an infinite gain limiter or relay whose input is a constant, X, plus a dither signal, D $\sin(\omega t - \varphi)$,

$$\text{input} = X + D \sin(\omega t - \varphi) \tag{A1}$$

In order to determine the intervals during which the input is positive or negative, consider sketch (x). Since the time origin is chosen so that



Sketch (x)

the input is zero for $t = 0$, equation (A1) can be solved to give

$$\varphi = \sin^{-1} \frac{X}{D} \tag{A2}$$

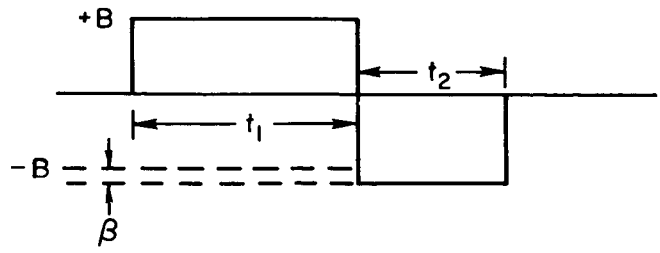
and from symmetry

$$t_1 = \frac{\pi}{\omega} + \frac{2\varphi}{\omega} \tag{A3}$$

also

$$t_2 = \frac{2\pi}{\omega} - t_1 = \frac{\pi}{\omega} - \frac{2\varphi}{\omega} \tag{A4}$$

the output is a rectangular wave as shown in sketch (y). The steady



Sketch (y)

component of this output is the average value, Y, of the rectangular wave which from geometrical relations is

$$Y = \frac{B(t_1 - t_2) - \beta t_2}{t_1 + t_2} \tag{A5}$$

or

$$Y = B\left(\frac{2\varphi}{\pi}\right) - \frac{\beta}{2} + \frac{\beta}{\pi}\varphi \tag{A6}$$

Substituting for φ from (A2) gives

$$Y = \frac{2}{\pi} B\left(1 + \frac{\beta}{2B}\right)\sin^{-1}\frac{X}{D} - \frac{\beta}{2} \tag{A7}$$

Notice that the average output given by equation (A7) consists of two parts: the first is proportional (for small values of X/D) to the input, X; the second is a constant or bias term which is a result of the unbalance of the limiter. This latter term shows that the successful use of such high-gain saturating systems depends to a large extent on having a well-balanced high-gain limiting device if zero steady state errors are to be attained. This effect of bias was also derived in reference 3 by a different means.

The low-frequency gain, Y/X, is given by (assuming $\beta = 0$)

$$\text{gain} = \frac{2B}{\pi X}\sin^{-1}\frac{X}{D} \tag{A8}$$

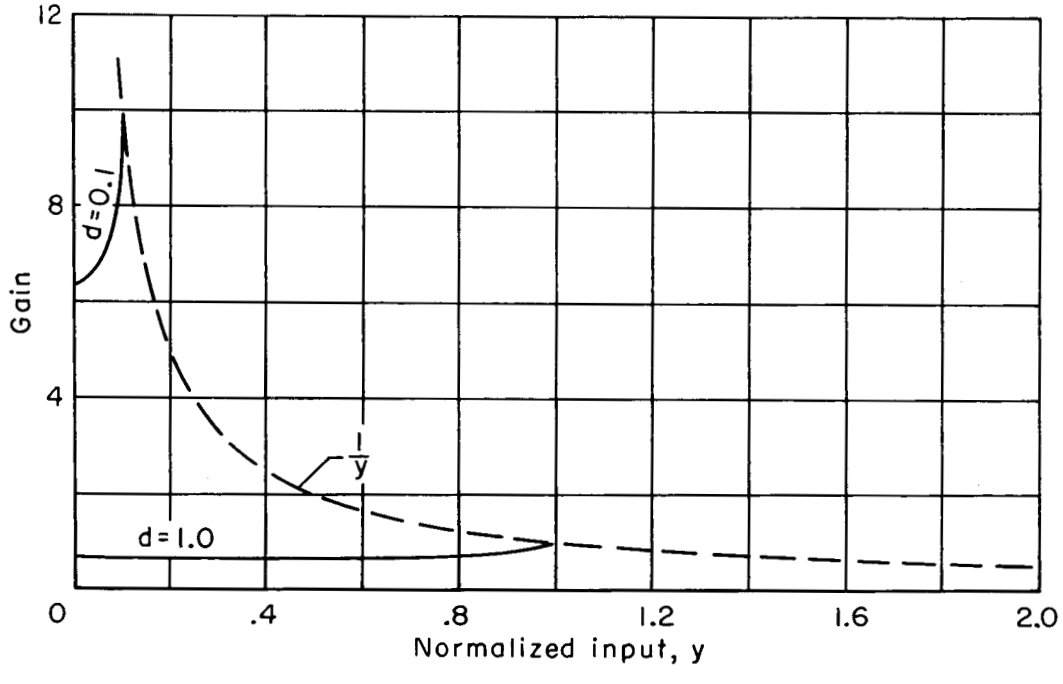
The expression can be normalized with respect to limit level by letting

$$\left.\begin{aligned} y &= \frac{X}{B} \\ d &= \frac{D}{B} \end{aligned}\right\} \qquad \text{(A9)}$$

which results in

$$\text{gain} = \frac{2}{\pi y} \sin^{-1} \frac{y}{d} \qquad \text{(A10)}$$

The gain as a function of the normalized input, y, is shown in sketch (z) for d = 0.1 and d = 1.0. The upper bound (dotted curve) is simply 1/y, the output (limit level) divided by the input X. As is illustrated in



Sketch (z)

this sketch, the gain is approximately constant for small values of y/d. The value of the constant gain is seen also to be dependent upon the normalized dither level, d.

The gain of the limiter for a sine wave input only (the dither) is, by describing-function analysis, the fundamental component of the output divided by the input amplitude. This gain is given by

$$\text{gain} = \frac{4}{\pi}\frac{B}{D} \qquad \text{(A11)}$$

Comparison of equation (A8) with equation (A11) shows that for small values of X/D, the gain for the dither is twice the gain for the constant input. This derivation shows why one may find the equivalent linear system by the simple means presented earlier in this report.

REFERENCES

1. Gregory, P. C., ed.: Proceedings of the Self Adaptive Flight Control Systems Symposium, Jan. 13-14, 1959. WADC Tech. Rep. 59-49, Mar. 1959.

2. Schmidt, Stanley F.: The Analysis and Design of Continuous and Sampled-Data Feedback Control Systems With a Saturation Type Nonlinearity. NASA TN D-20, 1959.

3. Flügge-Lotz, Irmgard, and Lindberg, Herbert E.: Studies of Second and Third Order Contactor Control Systems. Tech. Rep. No. 114, Division of Eng. Mech., Stanford Univ., 1958.

4. Flügge-Lotz, Irmgard, and Ishikawa, Tomo: Investigation of Third Order Contactor Control Systems, Part I. Tech. Rep. No. 116, Division of Eng. Mech., Stanford Univ., 1959.

5. Truxal, John G.: Automatic Feedback Control System Synthesis. McGraw-Hill Book Co., Inc., 1955.

6. Tsien, H. S.: Engineering Cybernetics. McGraw-Hill Book Co., Inc., 1954.

7. Van Meter, J. T., and McLane, C.: A Study to Determine an Automatic Flight Control Configuration to Provide a Stability Augmentation Capability for a High-Performance Supersonic Aircraft. WADC Tech. Rep. 57-349. (A series of four progress reports covering work done from Mar. 1, 1957 through Feb. 28, 1958)

TABLE I.- FORMULAS FOR ESTIMATING CHATTER FREQUENCY

System order	Approximate chatter frequency
4	∞
5	ω_1
6	$\omega_1\sqrt{\frac{P_1}{P_1 + 2\zeta_1\omega_1}}$
7	$\left[\frac{\omega_1\omega_2(b - \sqrt{b^2 - 4})}{2}\right]^{1/2}$ where $b = \frac{\omega_1}{\omega_2} + \frac{\omega_2}{\omega_1} + 4\zeta_1\zeta_2$

TABLE II.- MISSILE TRANSFER FUNCTION COEFFICIENTS

Flight condition	V	K_a	ω_a	ζ_a	$\tau_{\dot{\theta}}$
1	3340	0.11	3.6	0	12
2	3871	1.95	15	.15	1.1
3	4402	8.8	25	.30	.31

TABLE III.- CHATTER FREQUENCIES COMPUTED BY THREE METHODS

Case	Flight condition	Example	Chatter frequency computed by -		
			Formula	Root locus	Simulation
1	1	Fixed zeros	56.6	52.23	46.3
2	2	Fixed zeros		57.49	56.0
3	3	Fixed zeros		69.13	67.3
4	2	Moving zeros		50.97	48.5
5	3	Moving zeros		44.78	44.5

TABLE IV.- CHATTER AMPLITUDES COMPUTED BY THREE METHODS

Case	Chatter amplitude, g, (peak to peak) based on ω_c computed by -		
	Formula	Root locus	Simulation
1	0.000248	0.000466	Not measurable
2	.1144	.1078	0.12
3	1.578	.710	.76
4	.1144	.170	.190
5	1.578	4.08	4.2

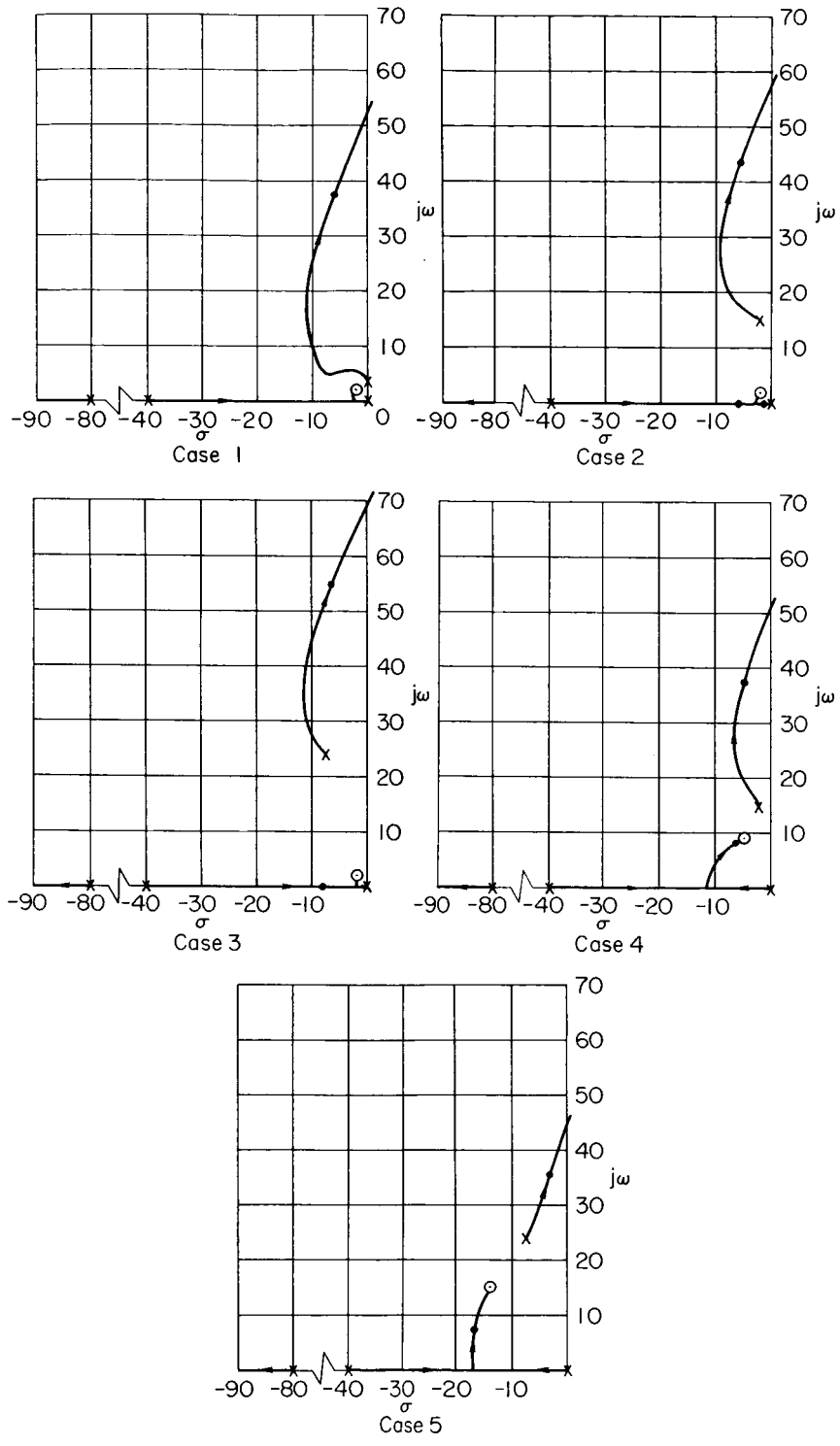


Figure 1.- Root locus graphs for the five cases simulated.

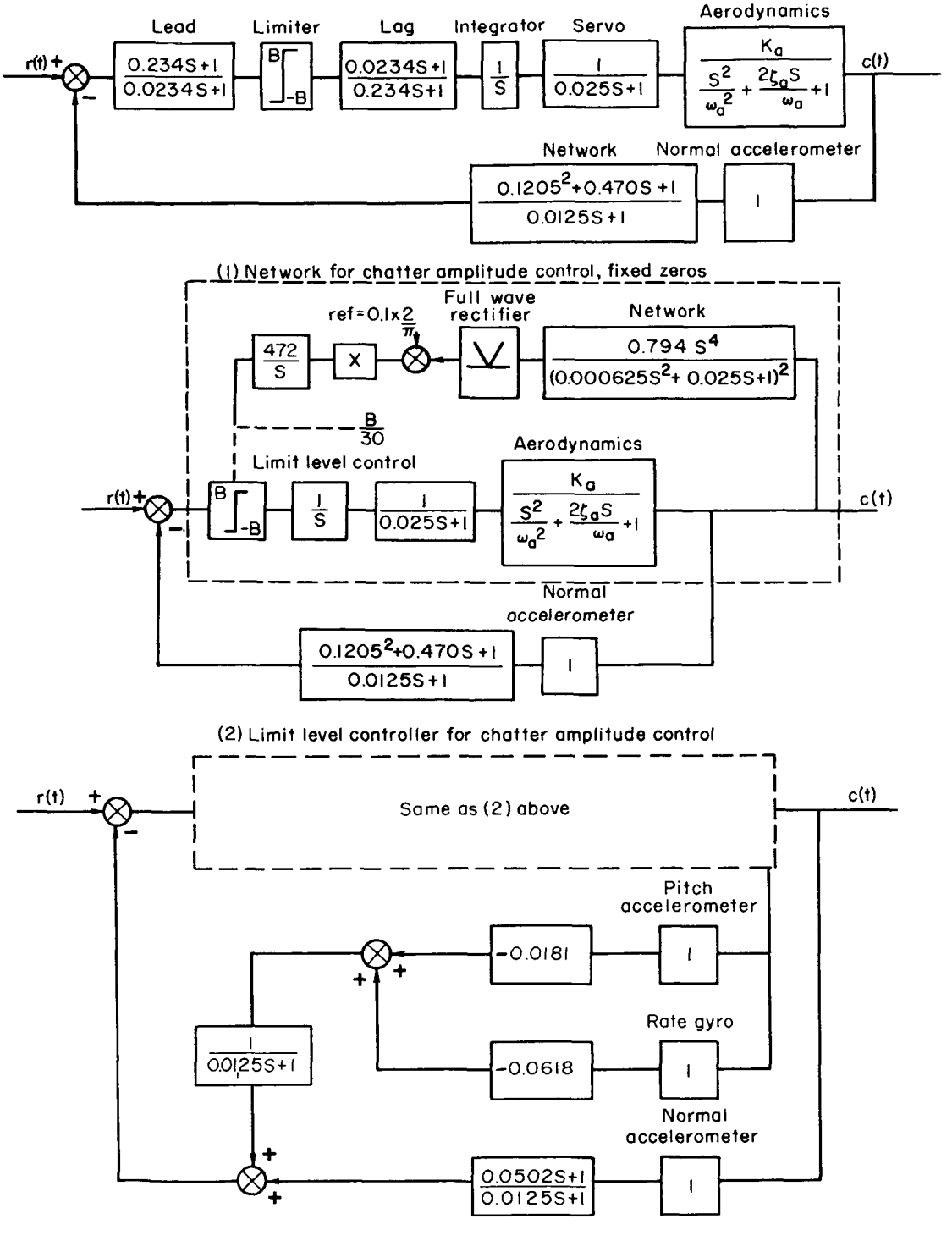


Figure 2.- Block diagrams of the three example systems.